Relativisation and Accessibility: A Corpus Analysis of Relative Clauses in Arabic Written Texts

Zainab Marzouk Al-Zaghir

A thesis submitted for the degree of Doctor of Philosophy in Linguistics,
University of Otago, Dunedin, New Zealand

August 2013

Abstract

Keenan and Comrie (1977) posited that the Noun Phrase Accessibility Hierarchy (NPAH) is a universal hierarchy determining the use of relative clauses in written and spoken communication. The NPAH states that the accessibility to relativisation follows the order, from the most to the least accessible: Subject (SU)> Direct Object (DO)> Indirect Object (IO)> Oblique (OBL)> Genitive (GEN)> Object of Comparison (OCOMP). Keenan (1975) suggested that the distribution of relative clauses in corpora reflects their accessibility, as predicted by the NPAH, with SU being the most accessible grammatical position to relativisation. Discourse functions of relative clauses, such as characterisation and anchoring, have also been argued to determine accessibility. Accordingly, Fox (1987) proposed the Absolutive Hypothesis (AH), which states that intransitive subject (ISU) relative clauses and (DO) relative clauses are more accessible than transitive subject (TSU) relative clauses. Other factors can influence the frequency of relative clauses in corpora, including genre (Jensen, 1999), spoken vs. written language (Gordon & Hendrick, 2005) and time period (Hogbin & Song, 2007). This thesis simultaneously considers the role of multiple factors on the frequency of relative clauses in Arabic texts.

Considered factors can be categorised into two types: (1) construction-external factors, which include language variety (dialect) and genre, and (2) construction-internal factors, which include characteristics of the head such as definiteness, grammatical position in the main clause and animacy. Among these factors, variety is of particular importance. One standard variety may not be enough to represent a language (Fleischer, 2004). Arabic countries commonly exhibit diglossia, where there are different varieties used (Ferguson, 1959). This study deals with relativisation in three varieties: Classical Arabic, Modern Standard Arabic and Iraqi Arabic. Thus, this study is one of the first to test a universal hierarchy on a dialectical basis.

In this study, relative clauses are classified into seven grammatical positions— TSU, ISU, DO, IO, OBL, GEN, OCOMP— and analysed statistically using multi-level Poisson regression analysis (Baayen, 2008; Keune, Ernestus, Van Hout, & Baayen, 2005). The results show that absolutive relative clauses (ISU+ DO) are the most frequent relative clauses used in Arabic texts, and this result shows a consistency

among the three varieties and the two genres (fiction and non-fiction). Moreover, IO relative clauses are not found to be as frequent as predicted by the NPAH.

The full order of relative clauses used in Arabic written texts is ABS> TSU, OBL> GEN> OTHERS, where OTHERS includes IO and OCOMP due to their small percentages. This order is found in Modern Standard Arabic and Iraqi Arabic, but not in Classical Arabic, where OBL is greater than TSU, and it is found in nonfiction but not in fiction where the TSU is greater than OBL. Furthermore, it is found that animacy of the heads of relative clauses plays a significant role in determining the distribution of relative clauses in the texts. Animacy determines whether the NPAH or the AH is manifested in the distribution of relative clauses in a text. The effect of animacy is not restricted to the upper grammatical positions, but also the lower positions, where the GEN position appears consistently higher than the OBL position when the heads of relative clauses refer to animate referents. Thus, accessibility to relativisation does not always follow a single universal hierarchy.

This thesis also investigates the use of relativisation strategies (the pronoun retention strategy and the gap strategy) in Arabic relative clauses. This study aims to find the basis on which these two relativisation strategies are used in light of the NPAH and other theories. The results show that the resumptive pronoun is mainly used to avoid ambiguous sentences and grammatical violations, in particular the preposition stranding rule. Moreover, there are several factors that determine the use of the resumptive pronoun over the gap including definiteness, the distance between the head and the relativised noun phrase, and the obligatoriness of the relativised noun phrase.

Acknowledgement

I would like to express my thanks to my first supervisor Professor Jae Jung Song for his guidance through my PhD. I wish also to express my gratitude to my second supervisor Dr. Hunter Hatfield, who helped me in statistics and continuously supported me until the very last minute before submission.

Special appreciation is extended to the University of Otago for awarding me with Otago scholarship; without this award, I would not be able to do my PhD. My thanks also go to all the members of the Department of English and Linguistics for their support. Special thanks go to Elizabeth Hogbin, who helped me in writing and proofreading my thesis.

I take the opportunity to record my thanks to Professor James Dickins (Professor of Arabic, University of Leeds). I am grateful to him for his valuable guidance in relation to the Arabic language.

I would also to express my sincere thanks to the library staff especially Thelma Fisher, who helped me a lot from the very early stages of my PhD in referencing using EndNote.

Last but not least, I would like to thank my husband, Atheer, and children, Qaswar, Jaafar and Durar for their unconditional love and support during the journey of my PhD. My thanks also go to my parents, who continued in supporting me and praying for me throughout my PhD. I am extremely grateful and indebted to all of you and I would not be able to finish my study without your encouragements and help.

Table of Contents

Abstract ..... **ii**
Acknowledgement ..... **iv**
List of Tables ..... **x**
List of Figures ..... **xi**
List of Abbreviations ..... **xiii**
Description of Arabic Sounds ..... **xiv**
List of Texts Used ..... **xvi**
Chapter One: Introduction **1**
1.1 Background to the Present Study 1
1.1.1 Related Studies 4
1.1.2 The Relativisation Strategies 8
1.2 This Study 10
1.2.1 The Focus of this Study 10
1.2.2 The Significance of the Study 12
1.2.3 Organisation of the Thesis 13
Chapter Two: Relativisation and Accessibility **16**
2.1 Introduction 16
2.2 The NPAH and Counter-Examples 17
2.2.1 The NPAH 17
2.2.2 Counter-Examples to the NPAH 18
2.3 Re-examining the NPAH 21
2.3.1 Subject Primacy 21
2.3.1.1 Theoretical Studies 22
2.3.1.2 Usage-Based Studies 26
2.3.1.2.1 Corpus Studies 26
2.3.1.2.2 Language Acquisition, Processing and Related Studies 35
2.3.2 The Use of Relativisation Strategies on a Continuous Segment of the NPAH 41
2.3.2.1 The Use of the RP and the NPAH 42

2.3.2.2 Pronoun Retention in Arabic 48
2.3.2.3 The Use of the RP and the Pragmatic Interpretation 54
2.4 Summary 56
2.5 Implications for This Study 58
Chapter Three: The Arabic Language 60
3.1 Introduction 60
3.2 Arabic Grammar 61
3.2.1 Parts of Speech: 61
3.2.1.1 The Verb 61
3.2.1.2 The Noun 65
3.2.1.3 Personal Pronouns 68
3.2.1.4 Prepositions 70
3.2.2 Simple Sentences: Structures and Patterns 72
3.2.3 Annexation 75
3.3 Relative Clauses 76
3.3.1 Types of Relative Clauses 77
3.3.1.1 Definite-Headed Relative Clauses 77
3.3.1.2 Indefinite-Headed Relative Clauses 78
3.3.1.3 Free/ Headless Relative Clauses 80
3.3.1.4 Restrictive and Non-restrictive Relative Clauses 81
3.3.2 Relativisation Strategies 82
3.3.2.1 The Resumptive Pronoun 82
3.4 Varieties of the Arabic Language 84
3.4.1 Relativisation in Arabic Dialects 86
3.5 Diglossia 87
3.6 Summary 89
Chapter Four: Methodology 90
4.1 Introduction 90
4.2 Selection of Texts 90
4.2.1 Varieties of Arabic Included in This Study 91
4.2.2 Genres 92
4.2.3 Other Issues 93
4.3 Relative Clauses 95

4.4 Test Methods....100
4.4.1 Grammatical Positions of the NPAH....100
4.4.1.1 Transitive vs. Intransitive Subject Relatives....102
4.4.1.2 The DO Position....108
4.4.1.3 The IO Position....110
4.4.1.4 The OBL Position....111
4.4.1.5 The GEN Position....112
4.4.1.6 The OCOMP Position....114
4.4.2 Statistical Analysis of the Data....115
4.4.3 The Pronoun Retention Strategy....118
4.5 Summary....120
Chapter Five: Results and Analysis....121
5.1 Introduction....121
5.2 Assessment of the NPAH and the AH....122
5.2.1 The NPAH....126
5.2.2 The AH....128
5.2.2.1 Functions of Arabic Relative Clauses....130
5.3 Definiteness....135
5.4 The Position of the Head in the Main Clause....140
5.5 Animacy....147
5.6 Variety....152
5.6.1 CA....155
5.6.2 MSA....159
5.6.3 IA....163
5.6.4 Summary....165
5.7 Genre....168
5.7.1 Fiction....170
5.7.2 Nonfiction....172
5.7.3 Summary....174
5.8 Summary....176
Chapter Six: Discussion....178
6.1 Introduction....178
6.2 General Results and Counts....178

6.3 The NPAH or the AH 183
6.3.1 The Results and the NPAH 183
6.3.2 The Results and the AH 185
6.4 The Role of Construction-External Factors 190
6.4.1Variety 190
6.4.1.1 Diglossia 192
6.4.2 Genre 193
6.5 The Role of Construction-Internal Factors 195
6.5.1 Definiteness 195
6.5.2 The Grammatical Position of the Head in the Main Clause 199
6.5.3 Animacy 201
6.6 Summary 204
Chapter Seven: The Pronoun Retention Strategy 206
7.1 Introduction 206
7.2 The NPAH 207
7.2.1 NPAH's Claims on the Arabic Language and Violations to These Claims 211
7.3 Factors that Affect the Use of RPs 214
7.3.1 The Distance between the Head and the Relativised Noun Phrase 214
7.3.2 Definiteness 216
7.3.3 Animacy of the Head 218
7.3.4 The Accessibility Theory 220
7.3.5 The RP as the Last Option 223
7.3.6 Studies on Arabic 225
7.3.7 Summary 233
7.4 The Interpretations of Relative Clauses with and without RPs 234
7.4.1 The Concept 235
7.4.2 RPs for Emphasis 235
7.4.3 The Restrictive Focus 237
7.5 The Use of the Pronoun Retention Strategy across Varieties of Arabic and genres 237
7.6 Summary 239
Chapter Eight: Conclusions 242
8.1 Introduction 242

8.2 Study Design....243
8.3 The General Distribution of Relative Clauses in Arabic Written Texts and Accessibility to Relativisation....243
8.3.1 Relativisation across Varieties....245
8.3.2 Relativisation in Fiction and Nonfiction....246
8.3.3 The Effect of the Construction-Internal Factors on the Frequency of Relative Clauses.....246
8.4 The Use of the Pronoun Retention Strategy in Arabic Relative Clauses 248
8.5 Limitations and Recommendations for Future Research....250
References....Error! Bookmark not defined.
Appendix A: Types of OBL and GEN Relative Clauses....262
Appendix B: Statistical Models Used in This Study....268
Appendix C: Classification of relative clauses according to the grammatical positions of the head NPs in the main clauses and in the relative clauses....293

List of Tables

Table 1: Inflection for Person, Number and Gender in the Past Form of the Verb.....62
Table 2: Inflection for Person, Number and Gender in the Present Form of the Verb ..63
Table 3: Inflection for Grammatical Cases in Arabic Singular Nouns 67
Table 4: Inflection for Number and Case in Arabic Non-Singular Nouns.....67
Table 5: Inflections in Free Personal Pronouns in Arabic.....69
Table 6: Inflections in Bound Personal Pronouns in Arabic 69
Table 7: The Relative Markers in Arabic 77
Table 8 : Texts Included in the Data Collection for this Study 94
Table 9: Model 1: Mixed Effect Poisson Regression for the Distribution of Relative Clauses in Arabic Texts.....123
Table 10: Model 2: Mixed Effect Poisson Regression for the Distribution of Relative Clauses Including Absolutive Relative Clauses 129
Table 11: Anchors in Arabic Relative Clauses.....131
Table 12: Model 3: Mixed Effect Poisson Regression for the Interaction between Different Anchors and Types of Relative Clauses 132
Table 13: Model 4: Mixed Effect Poisson Regression for the Interaction between Definiteness and Distribution of Relative Clauses in Arabic Texts 137
Table 14: The Distribution of Heads NPs among Grammatical Positions in the Main and Relative Clauses.....143
Table 15: NPs in Different Grammatical Positions and Relative Clauses Identifying Them.....143
Table 16: Model 5: The Interaction Between the Grammatical Position of the Head in the Main Clause and its Definiteness 145
Table 17: Model 6: The Interaction between Animacy and Relative Clauses in the NPAH 149
Table 18: Model 7: Mixed Effect Poisson Regression for the Interaction between Relative Clauses and Varieties 153
Table 19: Positions of Heads in Relative Clauses and in Main Clauses 159
Table 20: Grammatical Positions of the Heads in the Relative Clauses and in the Main Clauses in MSA Texts.....162
Table 21: Grammatical Positions of Heads in Relative Clauses and in Main Clauses in IA Texts.....165
Table 22: Model 8: Mixed Effect Poisson Regression for the Interaction between Relative Clauses and Genres 168
Table 23: Grammatical Positions of Heads in the Relative Clause and in the Main Clause in Fiction Texts.....172
Table 24: Grammatical Positions of Heads in the Relative Clause and in the Main Clause in Nonfiction Texts.....174
Table 25: Violations to the Predicted Distribution of RPs in Relative Clauses in Arabic Texts 211

List of Figures

Figure 1: A tree diagram of a subject-verb-object English sentence.....23
Figure 2: Distribution of Arabic relative clauses in written texts 122
Figure 3: The distribution of relative clauses in Arabic texts 125
Figure 4: The distribution of relative clauses in Arabic texts when TSU and ISU are combined into SU.....126
Figure 5: The distribution of OBL relative clauses according to their types 127
Figure 6: The distribution of GEN relative clauses according to their semantic relationships.....128
Figure 7: The interaction between anchors and relative clause types 133
Figure 8: The distribution of definite and indefinite relative clauses in Arabic texts ..136
Figure 9: The interaction between the grammatical position of the relativised noun phrase and definiteness.....139
Figure 10: The distribution of relative clauses according to the grammatical position of the head.....141
Figure 11: The classification of the category of 'OTHERS' according to their percentages within this category 141
Figure 12: The interaction between the grammatical position of the head in the main clause and its definiteness 147
Figure 13: The distribution of relative clauses whose heads refer to animate and inanimate referents 149
Figure 14: The interaction between relative clauses in different grammatical positions on the NPAH and animacy.....151
Figure 15: The interaction between relative clauses and varieties.....154
Figure 16: The distribution of CA relative clauses in fiction and nonfiction texts.....157
Figure 17: The classification of CA relative clauses according to their head's position in the main clause 158
Figure 18: The distribution of MSA relative clauses in fiction and nonfiction texts...161
Figure 19: The classification of CA relative clauses according to the grammatical positions of their heads.....162
Figure 20: The classification of IA relative clauses according to the grammatical positions of their heads.....164
Figure 21: The distribution of relative clauses in the three varieties of Arabic 166
Figure 22: The distribution of relative clauses that refer to animate and inanimate referents in the three varieties of Arabic 167
Figure 23: The interaction between relative clauses and genre in Arabic texts 169
Figure 24: The classification of relative clauses in fiction texts according to the grammatical positions of their heads in the main clauses 171
Figure 25: The classification of relative clauses in nonfiction texts according to the positions of their heads.....174
Figure 26: The distribution of relative clauses in fiction and nonfiction texts.....175
Figure 27: The percentage of relative clauses in each sentence type that includes RPs 207

Figure 28: The percentage of relative clauses using the RPs in relation to the distance between the head and the relativised noun phrase.....216
Figure 29: A comparison of the percentages of (def.) definite and (ind.) indefinite relative clauses with and without RPs.....217
Figure 30: A comparison of the percentages of inanimate, human and nonhuman animate relative clauses with and without RPs.....219
Figure 31: A comparison of the percentages of inanimate, human and nonhuman animate relative clauses, other than SU relative clauses, with and without RPs.....220
Figure 32: A comparison of the percentages of relative clauses whose scores are from zero - five with and without RPs.....223
Figure 33: Percentages of pronouns vs. full noun phrases that function as SU in DO relative clauses which use the gap strategy.....230
Figure 34: A comparison between the percentages of relative clauses with and without RPs when the subject position inside the relative clause is occupied by pronouns and full noun phrases.....231
Figure 35: A comparison of the percentages of relative clauses that use RPs and relative clauses that use gaps across the three varieties of Arabic.....238
Figure 36: A comparison of the percentages of relative clauses that uses RPs and relative clauses that use gaps in fiction and nonfiction.....239

List of Abbreviations

ABS: Absolutive
ACC: Accusative
ADV: Adverb
Ani: Animate
AH: Absolutive Hypothesis
CACC: Cognate Accusative
CA: Classical Arabic
DEF: Definite
DO: Direct Object
DU: Dual
EXIS: Existential
FEM: Feminine
FIC: Fiction
FUT: Future
GEN: Genitive
GENR: Genre
INANIM: Inanimate
IND: Indefinite
ISU: Intransitive Subject
IO: Indirect Object
IA: Iraqi Arabic
L1: First language
MAS: Masculine
MSA: Modern Standard Arabic
NOM: Nominative
NONFIC: Nonfiction
NP: Noun Phrase
NPAH: Noun Phrase Accessibility Hierarchy
OBL: Oblique
OCOM: Object Complement
OCOMP: Object of Comparison
P: Patient
PASS: Passive
PP: Prepositional phrase
SUCOM: Subject Complement
PRES: Present
PRO: Pronoun
R: Recipient
REL: Relative Marker
SG: Singular
SU: Subject
SPH: Subject Primacy Hypothesis
T: Theme
TSU: Transitive Subject

Description of Arabic Sounds

The transliteration used in this thesis follows the transliteration system used in Encyclopaedia of Arabic Language and Linguistics EALL.

Consonants

ʾ	voiceless glottal stop (no equivalent in English)
b	voiced bilabial stop, similar to /b/ in English
t	voiceless alveolar stop, similar to /t/
t̲	voiced interdental fricative, similar to /θ/
ǧ	voiced palate-alveolar fricative, similar to /ʤ/
ḥ	voiceless pharyngeal fricative (no equivalent in English)
ḫ	voiceless velar fricative (no equivalent in English)
d	voiced alveolar stop, similar to /d/
d̲	voiced interdental fricative, similar to / ð /
r	voiced alveolar trill, similar to Scottish /r/
z	voiced alveolar fricative, similar to /z/
s	voiceless alveolar fricative, similar to /s/
š	voiceless palatal fricative, similar to /ʃ/
ṣ	voiceless velarised alveolar fricative (no equivalent in English)
ḍ	voiced velarised alveolar stop (no equivalent in English)
ṭ	voiceless velarised alveolar stop (no equivalent in English)
ẓ	voiced velarised interdental fricative (no equivalent in English)
ʿ	voiced pharyngeal fricative (no equivalent in English)
ġ	voiced velar fricative (no equivalent in English, but similar to French /r/)
f	voiceless labiodental fricative, similar to /f/
q	voiceless uvular stop (no equivalent in English)
k	voiceless velar stop, similar to /k/
l	voiced alveolar lateral, similar to /l/
m	voiced bilabial nasal, similar to /m/
n	voiced alveolar dental, similar to /n/

h	voiceless glottal fricative, similar to /h/
w, ū	voiced bilabial semivowel, similar to /w/
y, ī	voiced palatal semivowel, similar to /j/
g	voiced velar stop, similar to /g/; it is found in IA but not in CA and MSA.
ch	voiceless palato-alveolar affricate, similar to /tʃ/; it is found in IA but not in CA and MSA.

Vowels

a	low short central unrounded vowel
ā	low long central unrounded vowel
i	high short front unrounded vowel
ī	high long front unrounded vowel
u	high short back rounded vowel

List of Texts Used

The following texts are used for the data collection of this study:

Al-‘ishbili, A. (1332-1406). *Tārīḫ Ibn ḫaldūn: Kitāb al-ʿibar wa diwān al-mubtadaʾ wa al-ḫabar fī tāriḫ al-ʿarab wa al-barbar wa man ʿāšarahum min ḏawi al-sulṭān al-ʾakbar ‘Ibn Khaldoun's history: Book of lessons, record of beginnings and events in the history of the Arabs and Berbers and their powerful contemporaries’*: International Ideas Home.

Alf laila wa laila ‘Arabian nights’. (1740). Beirut: Al-maktabah Al-shaʿbiah.

Al-hariri, A. A. (1873). *Maqāmāt Al-ḥarīrī.* ‘Al-Hriri’s maqamat’. Beirut: Matbaʿat Al-maʿaarif. (d. 1008)

Al-samarraaʾi, K., Taaha, A.-W., & Matloub, N. (2000). *Tārīḫ al-Aarab wa haḍāratihum fī Al-andalus ‘History of Arabs and their civilizations in Andalus’*. Beirut: Dar Al-kitab Al-jadīd.

Al-shaybani, I. (1231). *Al-kāmil fī al-tārīḫ: Tarīḫ Ibn Al-aṯīr ‘ The complete history: Ibn Al-atheer’s history’*: International Ideas Home.

Al-shaykh, R. (1996). *Tarīḫ al-Aarab al-muʾāsir ‘Contemporary Arab history’*. Cairo: Ein for Human and Social Studies.

Al-takarli, F. (1980). *Al-raǧʿ al-baʾīd ‘The far echo’*. Beirut: Dar Al-mada Lilthaqafah wa Al-finūn.

Al-tibari, M. (838-923). *Tarīḫ al-umam wa al-mulūk tarīḫ Al-ṭibarī ‘History of nations and kings Al-tibari's history’*: International Ideas Home.

Al-wardaani, S. (1999). *Al-saif wa al-siyasah fī al-Islam: al-sirāʿ bain al-Islam al-Nabawī wa al-Islam al-Amawī ‘The sword and politics in Islam: The conflict between the Prophet’s Islam and the Amawi’s Islam’* Beirut: Dar Al-Raʾi.

Al-yasiri, S. (1972). *Rubāʿiyāt Abu Gāṭiʿ ‘Abu Gati’s Quadruplet’*. Baghdad: Maṭbaʿat Al-shaʿb.

Badr, A. (2005). *Al-ṭarīq ʾilā tall al-muṭrān ‘The road to tal al-mutran’*. Beirut: Riad Al-Rayyes Books.

Furmaan, G. (1988). *Al-naḫlah wa al-ǧīrān ‘The palm tree and neighbours’*. Baghdad: Dar Al-farabi/ Dar Babil.

Sa'īd, M. (2008). *Taġrīdat al-baǧaʿah ‘The tweet of the swan’*. Beirut: Dar Al-adaab

Tufail, A. (1900). *Hayy Ibn Yaqḍān ‘Hayy, the son of Yagdhaan’.* Algeria: Fontanah wa Shraka'eh. (Original work published 1185)

Zaayid, A. (2006). *Al-manbūḏ ‘The castaway’*. Beirut: Al-daar Al-Aarabia lilʾulūm-Nashirūn.

Chapter One: Introduction

The fundamental importance of communication on the one hand and the diversity of languages and cultures in the world on the other hand raises the question, "in what ways do natural languages differ and in what ways are they all alike? "(Perlmutter, 1980, p. 195). Answering this question is the main aim of linguistic typology, which "concerns itself with the study of structural differences and similarities between languages" (Velupillai, 2012, p. 15).

Typologists have focused their attention on studying variation across languages, where the standard varieties are chosen to represent these languages. However, some grammatical features are found in dialects but not in standard varieties (Kortmann, 2004). That gap brings forth the need for a new trend in linguistics where typology can benefit from dialectology and dialectological studies can be performed using typological approaches.

The goal of this study is twofold. First, the study investigates accessibility to relativisation across varieties of the Arabic language. Second, it examines the use of the pronoun retention strategy in Arabic relative clauses. The way to achieve that goal and the objectives which underlie it will be outlined in this chapter.

1.1 Background to the Present Study

Relative clauses are clauses whose function is to modify an NP (Velupillai, 2012, p. 323). More precisely, a relative clause can be defined as " a clause narrowing the potential reference of a referring expression by restricting the reference to those

referents of which a particular proposition is true" (Comrie & Kuteva, 2011). There are two types of relative clauses: restrictive and non-restrictive. The focus of this study is restrictive relative clauses which are considered to be more common than non restrictive relative clauses (Givón, 1993).

Restrictive relative clauses "are used in grounding the referent noun to the knowledge-base that is already represented in the mind of the hearer" (Givón, 1993, p. 108), so they have an important role in effective communication in everyday life. Due to the importance of relative clauses in communication as well as the complexity of their structures (Diessel, 2009), they have been the focus of different synchronic and diachronic studies (Gómez, 2006).

"The relative clause [...] occupies a very prominent place in linguistic typology" (Song, 2001, p. 211). One of the significant cross-linguistic studies that investigate relativisation in different languages is Keenan and Comrie (1977), described as "one of the most influential works in the language universals literature" (Fox, 1987, p. 856). Based on the data of around fifty languages, Keenan and Comrie (1977, 1979) state that some grammatical positions are more accessible to relativisation than others, and that accessibility to relativisation follows an implicational hierarchy, the Noun Phrase Accessibility Hierarchy (NPAH)[1], which is as follows:

SU > DO> IO> OBL> GEN> OCOMP, where (>) indicates more accessible.

The following English examples are provided to demonstrate the grammatical positions of the NPAH:

SU (subject relative clauses), e.g. the man who bought the book...

DO (direct object relative clause), e.g. the man whom I met...

[1] Other abbreviations used in previous studies to refer to the Accessibility Hierarchy is the AH in Song (2001) and NP accessibility hierarchy in Croft (2003)

IO (indirect object relative clause), e.g. the man whom I gave the book to...

OBL (oblique relative clause), e.g. the house which I live in...

GEN (genitive relative clause), e.g. the man whose car is red...

OCOMP (object of comparison), e.g. the man whom I am taller than...

Keenan and Comrie presented the NPAH first in 1972, and then, in 1977, they presented the NPAH with a full account of methodological problems and counter examples. According to the NPAH, if a language can relativise only one grammatical position, then that position must be the SU position because it is the most accessible position on the NPAH. Moving down the NPAH, the difficulty of relativisation increases: IO, for example, is more difficult to relativise than DO, and OBL is more difficult to relativise than IO and so on. The OCOMP is the most difficult position to relativise. For example, according to the NPAH, to identify a man, the speaker will rather use SU relative clause "the man who bought the book" than IO relative clause "the man whom I gave the book" or the OCOMP relative clause "the man whom I am taller than".

Some languages can relativise all positions in the NPAH such as the English language; other languages can relativise only one grammatical position, the SU position, such as Malagasy, which relativises the other grammatical positions by promoting them into SU position through passivisation (Keenan & Comrie, 1977). As will be shown in chapter two, Arabic, is one of the few languages that can distinguish and relativise all grammatical positions on the NPAH.

The NPAH has been tested widely from a theoretical perspective (Hawkins, 1994; Johnson, 1977; Maxwell, 1979) and from a language acquisition perspective (Gas, 1979; Hyltenstam, 1984; Pavesi, 1986, among others). In spite of the counter

evidence to the NPAH which has been found in previous studies, it is considered as "broadly valid"(Croft, 2003, p. 150)

1.1.1 Related Studies

In an attempt to validate the NPAH's predictions in corpora, linguists have tested the distribution of relative clauses in written texts and spoken discourse against the NPAH. Keenan (1975) found that the frequency of relative clauses in English written texts conforms to the order of the NPAH. That is, SU relative clauses are used more frequently than DO relative clauses, which are used more frequently than IO and OBL relative clauses, and so on according to the order of the NPAH. Moreover, the ratio of SU to DO relative clauses correlates with the complexity of the text. The greater the difference between SU and DO relative clauses is, the easier the text is to comprehend, which indicates that SU relative clauses are used more frequently in easier texts; on the other hand, the smaller the difference between SU and DO relative clauses, the more difficult the text is to comprehend. The explanation given for these results is that the NPAH represents the psychological accessibility of relative clauses in different grammatical positions. Two conclusions among others can be drawn from Keenan and Comrie's (1975) study: First, the NPAH, which is founded on a cross-linguistic basis, "can determine performance constraints within languages"(Keenan, 1975, p. 147). Second, the SU position is psychologically more accessible to relativisation than the other grammatical positions on the NPAH; accordingly, SU relative clauses are more accessible to comprehension, acquisition and production than other relative clauses. However, the second conclusion has been challenged by Fox(1987), who proposed the Absolutive Hypothesis (AH) instead.

Fox (1987), on the basis of the data from English discourse, suggests that accessibility to relativisation in discourse is not determined by the grammatical position of the head in the relative clause, rather it is determined by the function of relative clauses in the text. Instead of what she calls 'the Subject Primacy hypothesis' which means the subject is more accessible than other grammatical positions to relativisation, she proposed the AH. The AH states that absolutive relative clauses: Intransitive Subjects (henceforth ISU, e.g. the man who looks handsome) and DOs (e.g. the man whom I met), are more accessible to relativisation than Transitive Subject (henceforth TSU, e.g. the man who bought the book) relative clauses, because of the discourse functions of the relative clauses "rather than a special cognitive status" (Fox, 1987, p. 869). Fox explains that

> "relative clauses serve to situate the referent that is being introduced as a relevant part of ongoing discourse; in a sense they justify the introduction of the referent in the first place"(Fox, 1987, p. 861).

Situating a referent in discourse can be achieved through two strategies. First, provide a static description of the referent; second, link or anchor[2] the referent into discourse through another referent which is well known to the addressee. The function of ISU relative clauses is to provide a description or characterization of the referent, for example, "she is married to this guy who is really quiet" (Fox, 1987, p. 859). On the other hand, the function of TSU and DO relative clauses is to anchor the referent into the text, as in "I know somebody who has **her** now" and "This man who **I** have for linguistics is really too much", respectively (Fox, 1987, p. 859). If both DO and TSU relative clauses have an anchoring function, then why does Fox suggest that DO

[2] Fox adapts the term 'anchor' from Prince (1981); "A discourse entity [= 'referent' in Fox's terminology] is Anchored if the NP representing it is LINKED, by means of another NP, or 'anchor', properly contained in it, to some other discourse entity"

relative clauses should be more predominant than TSU relative clauses? The answer is that the DO relative clause links or anchors the head of the relative clause into the context using the SU in the relative clause, as is shown in the example above where the anchor is 'I' (in bold type). The TSU relative clause, on the other hand, links the head into the context using the DO of the relative clause, which is 'her' in Fox's example. Noun phrases in the SU position mostly carry given information and tend to be pronominal, so they perform the anchoring function better than noun phrases in the DO position, which usually carry new information.

The AH explains the accessibility to relativisation in three grammatical positions only, TSU, ISU and DO; it does not account for the other grammatical positions on the NPAH. Thus, the question to be answered is "whether or not Fox's (1987) discourse-based explanation can also extend to the remaining lower portion of the [NP]AH" (Song, 2001, p. 241); this is one of the questions that this study addresses.

It might be argued that the difference between the results obtained by Keenan (1975) and Fox (1987) is attributable to the difference between natural spoken conversations, which is the corpus used by Fox, and written texts, which is the corpus used by Keenan. However, this argument has not been supported by the data of the previous studies on written texts such as Jensen (1999) and Hogbin and Song (2007); and studies on written and spoken corpora such as Gordon and Hendrick (2005), where the AH is found to be reflected in written and spoken discourse. In particular, Jensen (1999) has stated that genre plays an important role in the distribution of relative clauses in written texts in a way supporting the NPAH or the AH. This study will consider relativisation in two genres, fiction and nonfiction.

The research on the role of the NPAH and on the discourse functions of relative clauses in predicting accessibility to relativisation has been based on the standard

varieties of languages, ignoring the differences that might exist between the standard variety and other dialects of a language. In fact, data collected from the standard variety have been found "fairly unrepresentative if compared to the overall picture"(Fleischer, 2004, p. 236). This is found to be true in German (Fleischer, 2004) and English (Kortmann, Herrmann, Pietsch, & Wagner, 2005). Therefore, this thesis will investigate relativisation in three varieties of the Arabic language, Classical Arabic (CA), Modern Standard Arabic (MSA) and Iraqi Arabic (IA), and compare the data of these varieties to the predictions of the NPAH and the AH.

There are several studies that investigate the role of the NPAH in predicting accessibility to relativisation in different languages. However, each of these studies focuses on one factor only, for example, Gordon and Hendrick (2005) have considered the role of written vs. spoken language, Fleischer (2004) has focused on the role of varieties of the language, and Hogbin and Song (2007) have considered chronological differences (i.e. between two periods of time). Furthermore, previous studies of the NPAH have not taken into account characteristics of the head of the relative clause such as definiteness, animacy and grammatical position in the main clause, in spite of the fact that these factors have been found to play an influential role in determining accessibility to relativisation (as will be discussed in 1.2.1). This study is an attempt to remedy this state of affairs by examining how multiple factors affect accessibility to relativisation in the Arabic language, and how these factors might interfere with the NPAH's predictions. These factors are classified into two groups: construction-internal factors and construction-external factors. Construction-internal factors are related to characteristics of individual relative clauses and the matrix clause in which they are embedded. In this study, construction-internal factors include (1) definiteness, (2) animacy and (3) the grammatical position of the head in the main clause. On the other

hand, construction-external factors are predictors of relative clauses above the individual clause, which are (1) variety and (2) genre.

1.1.2 The Relativisation Strategies

There are four types of relativisation strategies that are used to form relative clauses cross-linguistically: gapping or obliteration; pronoun retention; relative-pronoun; and non-reduction; see Comrie (1989, pp. 147-153) and Song (2001, pp. 216-220) for detailed discussion of relativisation strategies. Some relativisation strategies reveal the role of the head in the relative clause, which Keenan and Comrie refer to as [+case]; an example of that is the pronoun retention strategy (1977, p. 65). The pronoun retention strategy is the "use of a personal pronoun in the restricting clause, which is co-referential with the head noun"(Song, 2001, p. 218). An example of the use of the pronoun retention strategy is the relative clause:

Al-rağul allaḏī ra'itu-***hu***
The-man REL(3.SG.MAS) saw-***him***
'*[3]The man that I met ***him***'

In this example *-hu* 'him' is a resumptive pronoun which is co-referential with the head noun phrase *al-rağul* 'the man'. Since 'him' occupies the DO position in the relative clause, this clearly indicates that 'the man' is the DO of the relative clause. Thus, the pronoun retention strategy is an explicit or [+case] strategy. On the other hand, other strategies such as the gapping strategy are considered as [-case] strategies because the role of the relativised noun phrase cannot be recovered from the surface structure of the relative clause, as in the example: 'the man that I met in the street'. In this example, there is no indication that informs the reader that 'the man' is the DO of the relative clause.

[3] '*' ,if used with sentences, indicates that the sentence is ungrammatical.

One of the NPAH's constraints is that relativisation strategies should be used in a continuous segment. That is, if the pronoun retention strategy is used in SU and IO positions then it should also be used in the DO position. A relativisation strategy is considered the primary strategy of the language if it is used to relativise the SU in that language. If the primary strategy is used in a certain position, it should be used in all the upper positions and it may "cut off at any point on the [NP]AH" (Keenan & Comrie, 1977, p. 68). In the Arabic language, the primary strategy, the gap strategy, is used in SU relative clauses only, whereas noun phrases in other grammatical positions can be relativised using the pronoun retention strategy (Keenan & Comrie, 1979, p. 333). This study will focus on the pronoun retention strategy and the gap strategy as used in Arabic.

Several studies have been conducted to examine the basis on which the distribution of the pronoun retention strategy vs. the gap strategy is found in languages. According to the NPAH, the pronoun retention strategy is used in positions down the NPAH more than upper positions because explicit strategies are used in difficult contexts where the meaning cannot be perceived otherwise (Keenan & Comrie, 1977, p. 92). Therefore, in Arabic, Keenan and Comrie claim that the pronoun retention strategy is obligatorily used in all positions of the NPAH except in SU relative clauses (1979, p. 333).

Some factors are found to be influential in the choice of the gap strategy over the pronoun retention strategy or vice versa; among these factors are the distance between the position of the relativised noun phrase and the head (Ariel, 1990; Collier, 1991; Prince, 1990); definiteness (Galal, 2004; Shlonsky, 1992; Suaieh, 1980) and animacy (Erteschik-Shir, 1992). Ariel (1990, 1999) suggests that the use of the pronoun retention strategy is better predicted when considering more than one factor at the same

time. These factors are: the distance between the head and the relativised noun phrase; the restrictiveness of the relative clause; the complexity of the head and the obligatoriness of the relativised noun phrase. On the other hand, several studies (e.g. Doron, 1982; Jassim, 2011; Sells, 1984, 1987) suggest that pragmatic factors such as focus can play an important role in the choice of the resumptive pronoun over the gap or vice versa. This study will build upon the results of the previous studies and test the previous claims on the use of pronoun retention strategy against Arabic data. Thus, this study will attempt to find out how syntactic and semantic factors affect the use of the pronoun retention strategy in Arabic relative clauses.

1.2 This Study

Based on Keenan and Comrie's (1977) definition of relative clauses, only restrictive relative clauses with heads will be included in the data of this study. According to previous studies, mainly Keenan (1975) and Fox (1987), accessibility to relativisation can be measured through the frequency of relative clauses in corpora. That is, the number of relative clauses provides an indication of which grammatical relations are relativised more frequently than others, and so they are considered more accessible than others. This study, similar to previous studies, uses the frequency of relative clauses to answer the research questions listed in 1.1.1.

1.2.1 The Focus of this Study

As has been mentioned earlier, the focus of this study is investigating relative clauses in three varieties of Arabic, CA, MSA and IA. Genre is also taken into consideration, so data will be collected from fiction and nonfiction texts. When analysing the data of this study, three factors will be taken into account, which have

been referred to previously in this chapter as construction-internal factors. The first of these factors is definiteness. Only relative clauses with definite heads are included in Keenan and Comrie (1977) and Keenan (1975), yet both relative clauses with definite heads and those with indefinite heads are included in Fox (1987) and other corpus studies. Similarly, in this study, both relative clauses with definite and indefinite heads are included. Definiteness will be considered a variable whose effect on the distribution of relative clauses will be measured statistically.

Considering definiteness as an influential factor also arises from the fact that, in the Arabic language, the difference between definite and indefinite relative clauses is not merely semantic (as will be discussed in chapter two). Definite relative clauses use relative markers while indefinite relative clauses do not. Since the presence of the relative markers affects the accessibility of relative clauses (Prideaux & Baker, 1986, p. 39), definiteness in Arabic is considered an important factor that should be taken into consideration when investigating accessibility to relativisation.

The second factor that is considered in analysing the data of this study is animacy. The tendency is different among grammatical positions of the NPAH as to whether they have animate heads. For example, the heads of SU relative clauses tend to be animate whereas the heads of DO relative clauses tend to be inanimate (Roland, Dick, & Elman, 2007). Furthermore, DO relative clauses which occur with animate heads, will be harder to process than DO relative clauses which have inanimate heads whereas SU relative clauses with animate heads are easier to process than SU relative clauses with inanimate heads (Fernandez-Duque, 2009; Mak, Vonk, & Schriefers, 2002). Thus, animacy also has an influential role in determining accessibility to relativisation.

The third factor is the grammatical position of the head in the main clause. Although Song (2001, p. 223) states, "the grammatical relation [position] of the head noun in main clause turns out to be typologically far less interesting", the results of some studies (e.g. Kuno, 1974; Sheldon, 1974; Slobin, 1971) suggest that this factor is an important factor that influences the accessibility of relative clauses. Sheldon (1974) suggests that when the grammatical position of the head in the relative clause parallels its position in the main clause, the relative clause will be easier to process; this hypothesis is called the 'parallel hypothesis'. Thus, according to the parallel hypothesis, SU relative clauses can be categorised into different classes according to the grammatical position of the head in the main clause such as, SU-SU, DO-SU, OBL-SU and so on. SU-SU relative clauses are more accessible than other relative clauses.

Furthermore, in the Arabic language, the case of the head and the relative marker, if it is inflected for case, is governed by its grammatical position in the main clause not its grammatical position in the relative clause. That is, if the head noun phrase occupies the DO position in the main clause, regardless of its grammatical position in the relative clause, it is marked by accusative case. This is referred to as 'case attraction', which is also found in the grammars of ancient Greek and Latin (Comrie, 1989, pp. 153-154). Thus, the grammatical position of the head in the relative clause might have an influential role in accessibility to relativisation in the Arabic language.

1.2.2 The Significance of the Study

This study is one of the first attempts to test a universal hierarchy, the NPAH, against data from different dialects. The integration between typology and dialectology by "judg[ing] the observable patterns of cross-dialectical variation for individual

grammatical phenomena against generalisations, hierarchies and explanations which have grown out of the study of cross-linguistic variation" (Kortmann, 2004, p. 1) is an important trend in linguistics. Dialectal studies are found to be important in enhancing typological studies; "typologists will get a broader and, most likely, more adequate picture of what a given language is like if they no longer ignore dialectical variation" (Kortmann et al., 2005, p. 1). For example, a comparison among varieties of the German language shows that there is a significant difference between the standard and non-standard varieties of German, where relative clauses in non-standard varieties have more particles and resumptive pronouns than relative clauses in Standard German (Fleischer, 2004, p. 211).

This study does not only consider the differences among varieties. It also takes into consideration other variables such as genre, definiteness, animacy, the grammatical position of the head in the main clause and the discourse functions of relative clauses. By considering all these variables, this study will first provide us with a broad picture of relativisation in Arabic; second, it will improve our understanding of the role of universal hierarchies, such as the NPAH, in predicting accessibility to relativisation and the distribution of relative clauses in corpora.

1.2.3 Organisation of the Thesis

This thesis is organised as follows:

Chapter Two will present a review of the related literature. It will summarise the studies conducted on the NPAH, including both theoretical and language-based studies. The focus of the chapter will be on corpus studies that investigate the validity of the NPAH and the AH. It will also provide an overview of the studies that have been

conducted to examine the use of the pronoun retention strategy in Arabic and other languages.

Chapter Three will provide an overview of the Arabic language. It presents the reader with information about the structure of Arabic sentences in general and Arabic relative clauses in particular. It will also provide an introduction to the relativisation strategies used in the Arabic language.

Chapter Four will provide a discussion of the methodology used in this study. It will describe the procedures used in selecting the texts and the process of data collection. It will also describe the statistical method that is used to analyse the data and how this method is found useful in the case of this study, where there are multiple factors that are considered.

Chapter Five will present the results obtained through counting relative clauses. The results are presented in relation to the NPAH and the AH. It will also present the results of the interaction between the frequency of relative clauses and the variables that are tested in this study, which are definiteness, animacy and the grammatical position of the head in the main clause, in order to examine how these variables might interfere with the NPAH's predictions. The overall results are presented first, and then the results of the distribution of relative clauses across the three varieties and the two genres are presented one by one.

Chapter Six will present a discussion of the results in light of the research questions. It will show whether the results support the NPAH or the AH. It will also present a discussion of how construction-external factors, i.e., variety and genre, and construction-internal factors, definiteness, animacy and the grammatical position of the head, affect the distribution of relative clauses in Arabic written texts.

Chapter Seven will present the results on the pronoun retention strategy in relation to the NPAH and other theories. It will provide a discussion of how different factors contribute to the use of the pronoun retention strategy. Among the factors, which are discussed in Chapter Seven, are the distance between the relativised noun phrase and the head, animacy, definiteness and the nature of the Arabic language, which prevents preposition stranding. A discussion of how the use of the resumptive pronoun affects the interpretation of relative clauses will also be presented in Chapter Seven.

Chapter Eight will present the conclusions of this study. It will provide concluding remarks on the distribution of relative clauses in Arabic written texts as regards variety and genre. It will also provide conclusions on the effect of the studied construction-external and construction-internal factors on the distribution of relative clauses in written texts. The conclusions on the use of the pronoun retention strategy in Arabic relative clauses will also be presented. Finally, Chapter Eight will also discuss the implications and applications of this study.

Chapter Two: Relativisation and Accessibility

2.1 Introduction

During the last four decades, there has been a continuing interest in the NPAH established by Keenan and Comrie (1977). Many studies have been conducted to reinvestigate the NPAH's constraints in different languages, yet not all of these studies have provided support for the NPAH. This study will build on previous studies by examining the NPAH's constraints across three varieties of the Arabic language, CA, MSA and IA, and in two genres, fiction and nonfiction. Before presenting the methods used in collecting and analysing the data and the results of obtained in this study, it is helpful to present and discuss the previous studies conducted on the NPAH.

This chapter will first provide a revision of the NPAH and its hypotheses followed by some counterexamples mentioned by Keenan and Comrie (1977). Then, studies that have been conducted to test the NPAH will be classified into two types: first, studies that test the NPAH's first hypothesis, the Subject Primacy hypothesis (theoretical studies will be reviewed in section 2.3.1.1 and studies conducted on a language-usage basis will be reviewed in section 2.3.1.2); second, theoretical and language-based studies which have been conducted to examine the second hypothesis of the NPAH, which is the use of relativisation strategies on a continuous segment of the NPAH (focusing on the pronoun retention strategy in 2.3.2.1, with specific reference to Arabic in 2.3.2.2 and to other languages in 2.3.2.3). Lastly, a summary of this chapter is presented in section 2.4, followed by implications for the current study in 2.5.

2.2 The NPAH and Counter-Examples

In this section, I will review the NPAH and then present the languages Keenan and Comrie mention as counterexamples to the NPAH.

2.2.1 The NPAH

As is explained in section 1.1, Keenan and Comrie (1977) argued that different languages follow one universal hierarchy in relativisation, the NPAH. One important hypothesis of the NPAH is that the SU is the most accessible grammatical relation. According to the NPAH, SU is the grammatical relation that all languages with relative clauses must be able to relativise. Moving down the NPAH, the difficulty of relativisation increases: IO is more difficult to relativise than DO, and OBL is more difficult to relativise than IO and so on. The OCOMP is the most difficult position to relativise. The hypothesis that SU is the most accessible position to relativise is referred to by Fox (1987) as the 'subject primacy hypothesis' and is discussed further in section 2.3.1.

The second important hypothesis is that the use of any relativisation strategy should apply to a continuous segment of the NPAH (i.e. it should be used continuously on the positions of the NPAH with no gaps). Languages may differ in the strategies by which they relativise the positions on the NPAH. Moreover, for each language, there is one primary strategy, which is the strategy that is used to relativise the SU position. Languages might have strategies other than the primary but do not need to. According to the NPAH, the primary relativisation strategy should be used over a continuous segment of the NPAH. Thus, if a strategy, e.g. the pronoun retention strategy, works for SU and OCOMP positions on the NPAH, but does not work similarly on any of the middle positions on the NPAH (i.e. DO, IO, OBL, GEN), then there is a gap which

breaks the continuity of the NPAH. Nonetheless, any strategy which is used to relativise SU relative clauses may not be used with any lower position on the NPAH; another strategy can be used instead. If the primary strategy is used in a low position, it should be used in the higher positions. The hypothesis of using relativisation strategies continuously on the NPAH is addressed in this study in section 2.3.2.

Keenan and Comrie (1977, 1979) have provided a description for the relativisation strategies used in the fifty languages included in their sample. However, the description that they have provided is broad, especially for languages that have more than one variety. They considered only one variety of each of the languages included in their sample. Including only the standard varieties of languages to represent these languages might affect the reliability of their results. This is particularly true of the Arabic language because of the diglossic nature of the language (sections 3.4 and 3.5).

Keenan and Comrie (1977, 1979) provided evidence for the NPAH, but there are some counter-examples that the authors discuss in their study. These counterexamples will be explained in the next section.

2.2.2 Counter-Examples to the NPAH

Although, in Keenan and Comrie (1977, 1979), the NPAH was supported by approximately fifty languages, there are a few languages that appear as counterexamples to the NPAH, such as Dyirbal, Hausa, and Yoruba. Dyirbal is considered as an ergative language because "full NP subjects of intransitive sentences and full NP DOs of transitive sentences are case-marked in the same way (zero, in this

case), whereas subjects of transitive sentences carry a special marker, the ergative" (Keenan & Comrie, 1977, p. 82).

In the relativisation process, Dyirbal can relativise Absolutives (ISU and DO) only. On the other hand, ergatives (transitive subjects) cannot be relativised unless they are promoted into Absolutives by antipassivising the sentences before relativising them. Therefore, Johnson (1974) suggests that in order for the NPAH to predict patterns of relativisation in ergative languages, it should be changed as follows:

ABS > ERG > IO > OBL > GEN> OCOMP.

However, Keenan and Comrie (1977) provide examples from other ergative languages such as Basque, Walbiri, and Hindi that support the NPAH. Therefore, Keenan and Comrie consider the case of Dyirbal to be an exception, and suggest it should be treated differently. Keenan and Comrie suggest an explanation which makes the Dyirbal language support the NPAH; this explanation is "to regard the absolutive NP in transitive sentences (as well as in intransitive ones) as the subject" (Keenan & Comrie, 1977, p. 83). They explain that the properties of absolutive NPs in Dyirbal such as being essential in the sentence, preceding the ergative NP in the sentence and other characteristics which are discussed in Keenan and Comrie (1977, p. 84) are a subject's characteristics.

Hausa, Yoruba and Tongan on the other hand provide counter evidence to the second hypothesis of the NPAH since they "appear to use RC (relative clauses) strategies that apply to discontinuous segments of the RC" (Keenan & Comrie, 1977, p. 85). These languages use a [+case] strategy, pronoun retention, on the SU position and OBL, but they use a [–case] strategy, pronoun deletion, on DO. In the IO position, it is optional to retain the pronoun or to use the gap strategy. To further illustrate the

discontinuous segment in these languages, I present examples of SU, DO and OBL relative clauses in Hausa, which are provided by Keenan and Comrie (1977, p. 85); the RPs appear in bold, as is shown in examples (1)-(3).

1. dokin da **ya** mutu
   horse Rel **it** died
   The horse that died

2. mutumin da na gani (***shi**)
   person Rel I saw **him**
   The person that I saw

3. wuqad da ya kasha ta da **ita**
   knife Rel he killed her with **it**
   The knife with which he killed her

As is demonstrated in example (2), the RP *shi* 'him' is unacceptable. On the other hand, the RPs *ya* 'it' and ita 'it' are essential in (1) and (3). Thus, there is a gap on the DO position which means relativisation strategies in Hausa do not apply to a continuous segment of the NPAH. Keenan and Comrie (1977) explain that pronouns are found beside subjects in simple sentence and not only in relative clauses, which makes the use of the pronoun a verb agreement strategy more than a pronoun retention strategy. The same explanation can be applied to Yoruba, but not to Tongan.

The gap that breaks the segment of the relativisation strategies in Tongan has been explained historically. Keenan and Comrie (1977, p. 87) state, "the ergative subject of Tongan is a passive agent, and the absolutive DO in Tongan is historically

the subject". They explain that although the ergative noun phrase is claimed to have several subject properties, it cannot be considered as subject because it can be omitted when unspecified, like the agent in the English passive sentence. Thus, what appeared to be counter-examples have been explained in a way which makes them conform to the NPAH.

2.3 Re-examining the NPAH

Many studies have been conducted to validate the NPAH. These studies have been conducted on different bases: first on a theoretical basis (e.g. Hawkins, 1994; Johnson, 1977; Maxwell, 1979), second on a language-usage basis (e.g. Eckman, Bell, & Nelson, 1988; Fox, 1987; Gas, 1979; Gordon & Hendrick, 2005; Hyltenstam, 1984; Keenan, 1975). Studies in the second type examine the NPAH's predictions either according to the use of relative clauses in a corpus (Fox, 1987; Gordon & Hendrick, 2005; Keenan, 1975) , or according to the comprehension and production of relative clauses by language learners (e.g. Gas, 1979; Hyltenstam, 1984; Pavesi, 1986). In general, these studies can be divided into two types: first, studies that examine the subject primacy hypothesis, which are presented in 2.3.1, second, studies that examine the use of relativisation strategies according to the NPAH with a focus on the pronoun retention strategy, which are presented in 2.3.2.

2.3.1 Subject Primacy

One major proposal of the NPAH is the Subject Primacy Hypothesis (henceforth SPH). The SU position is easier to relativise than other grammatical relations, and SU relative clauses are easier to process, comprehend, acquire and

produce than other relative clauses. This claim should be valid across languages as well as within languages. This section will review the studies that examine the NPAH focusing on the subject primacy hypothesis. Section 2.3.1.1 reviews theoretical studies and section 2.3.1.2 reviews language-use based studies. Section 2.3.1.2 includes corpus studies and language acquisition studies in 2.3.1.2.1 and 2.3.1.2.2, respectively.

2.3.1.1 Theoretical Studies

There are a number of theoretical studies which examine the NPAH proposed by Keenan and Comrie (1977) and its hypotheses, in particular the SPH. Among these studies is Hawkins (1994, 2004), in which the theoretical basis of the NPAH is examined. Using a processing-based theory of relativisation, Hawkins (1994, 2004) provides a potential explanation and support for the NPAH. Hawkins' theory states that:

> "the human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or dependency are processed" (Hawkins, 2004, p. 31).

Hawkins' theory indicates that the more nodes (grammatical relations) that can be minimised in certain structures, the more these structures are preferred. According to Hawkins, "The structural complexity of an SD is measured by counting the set of nodes within it: the complexity of SD_i (structural domain $_{i)}$ exceeds that of SD_j iff $SD_i > SD_j$" (Hawkins, 1994, p. 29). For example, considering the tree of subject- verb- object in English, while the subject has two nodes [S and VP], the DO has four nodes: [S, NP(SU), VP and V]; which makes

the subject a part of the structural domain of the object and not vice versa, as is shown in Figure 1, adapted from Hawkins (1994, p. 27).

Applying this theoretical construct to relative clauses shows that the minimum number of nodes that represents the SU relative clause on a tree diagram is lower than the number of nodes that represents the DO relative clause. For example, the number of nodes in the SU relative '*the boy who runs*' is four, while the number of nodes in the DO relative '*the car which the boy bought*' is five. Therefore, SU is easier to relativise than DO. The SU position is the easiest one to relativise on the NPAH since it has the least number of nodes that are required for the comprehension of the relative clause; nodes are added on each position going down the NPAH. As a result, the OCOMP is the most difficult position on the NPAH as it has the highest number of nodes among all the positions on the NPAH. Hawkins' study provides a good explanation and support for the NPAH and its hypothesis 'the subject primacy'. However, an application of this explanation in VSO and OSV languages and in languages where heads of relative clauses are internal and final could be necessary to better support the idea of universality carried out by the NPAH.

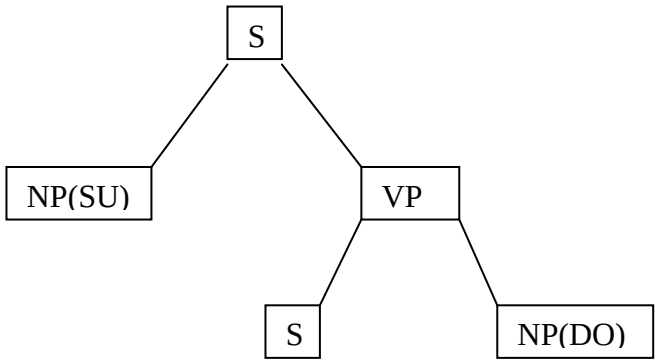


Figure 1: A tree diagram of a subject-verb-object English sentence

Considering the constituents of different constructions, Sigurd (1989) suggests that the NPAH, which is based on the grammatical role of the noun phrase in the relative clause, does not cover different structures or types of the same category. For example, Sigurd distinguishes between pied-piping, as in '*the boy to whom I talked*', and preposition stranding structures such as '*the boy whom I talked to*' which might be both considered to be OBL relative clauses; and he also distinguishes between the s-genitives and *of*-genitives. Sigurd suggests another hierarchy which he calls the 'supplementary hierarchy', "where the positions of the relativised element in the constituency trees of the relative phrases introducing the relative clause will play major roles" (Sigurd, 1989, p. 36). Accordingly, the s-genitive such as '*the boy whose car is stolen*' is more accessible than a structure such as 'the boy of whom the car is stolen', and relative clauses where preposition stranding is used are more accessible than relative clauses with pied-piping structures. Moreover, Sigurd suggests that the NPAH predicts the accessibility of simple relativised noun phrases whereas the supplementary hierarchy predicts the accessibility of complex relativised noun phrases, as in "*the boy to whose father the dog came*" and "*the boy to the father of whom the dog came*" (1989, p. 44). To supplement the NPAH, which is founded on a cross-linguistic basis, Sigurd's proposal needs to find support cross-linguistically, and that will be difficult as there is only one form of OBL and GEN in some languages, such as Arabic.

Other studies suggest that relativisation is influenced by factors such as semantics, pragmatics, context and language-specific syntactic rules. Among these studies is Matsumoto (1990) which suggests that the NPAH "does not seem suited to determine the preferred reading of a construction" in Japanese (p. 119), as in example (4), which is provided by Matsumoto (1990, p. 115), where the Japanese relative clauses have three interpretations:

4. hon o katta gakusei
   book ACC bought student
   The student who bought a book
   The student from whom Ø[4] bought a book
   The student for whom Ø bought a book

Matsumoto suggests that the acceptance and interpretation of relative clauses can be determined by different pragmatic factors such as the context and the hearer's view of the world. Since, in Japanese, acceptance and interpretation of the relative clauses depends on the reader/ hearer rather than on the grammar of the language, as there are no case markers to show the grammatical positions of the relativised noun phrases, theoretical analysis only might not provide helpful results; conducting an experiment such as a judgement experiment would reveal a more accurate result regarding the predictions of the NPAH. Similarly, Cole and Hermon (2005) suggest that the "appearance of conformity (of Indonesian relative clauses) to the (NP)AH is epiphenomenal" (p. 83). The authors explain that relativisation in the Indonesian language is determined by the syntax of the language, as in the case of the ungrammaticality of OBL relative clauses, which is due to the prohibition of preposition stranding. The syntax of the language might determine the relativisation strategy used to relativise a grammatical relation, yet it does not explain why some relative clauses are used more than others. For example, the syntax of the Indonesian language does not explain why SU relative clauses "occur much more frequently than non-subject relatives" (Cole & Hermon, 2005, p. 60), which is predicted by the NPAH.

This section has presented theoretical studies that have been conducted to re-examine the NPAH and Keenan and Comrie's (1977) study. Some of these studies have

[4] Ø indicates that there is a missing argument as "there is no rule requiring all the arguments of a verbal to be present in a sentence"(Matsumoto, 1990, p. 113)

examined the theoretical concerns of the NPAH universally, whereas other studies have tested the NPAH's predictions theoretically on individual languages. The next section will review some studies that have also been conducted to re-examine the NPAH, but on a language-usage basis.

2.3.1.2 Usage-Based Studies

In addition to the studies in the previous section that examine the predictions of the NPAH theoretically, there are some usage-based studies (corpus and language acquisition studies) which have been carried out to validate the NPAH.

2.3.1.2.1 Corpus Studies

Several corpus studies have been conducted to examine the NPAH's predictions. The first study is Keenan (1975) which was conducted to verify the validity of the NPAH by analysing data from written English texts. The author suggests that the NPAH "serves not only as the basis of universal constraints on the grammars of languages, but can determine performance constraints within languages"(Keenan, 1975, p. 147). In Keenan's study, 2200 relative clauses were collected from different kinds of written texts. Based on the NPAH, the author hypothesized three 'performance predictions'. First, in written texts, SU relative clauses should be used more often than DO relatives, which in turn should be used more often than IO relatives, and so on, following the order of the NPAH. This prediction has been partially confirmed since the percentage the author found for SU is (46.16%) followed by DO (23.73%), then OBL+IO (14.92%), the GEN (5.00%), and finally OCOMP (0%). In fact, Keenan

grouped IO with OBL and, therefore, did not provide evidence supporting IO by itself through his corpus study.

Second, SU relatives should be produced in simple texts more often than they are produced in complex ones. Therefore, the more complex the text is, the lower the ratio of subject relative clauses to object relative clauses is and vice versa. The number of DO and other relative clauses down the NPAH are used more frequently in complex texts than in simple ones so the difference between DO and SU relative clauses in complex texts is less than the difference between them in simple texts. This prediction is also supported, as the results show that in complex texts, the percentage of SUs is (40.12 %) followed by DOs (26.21%), while in simple texts the percentage of SUs is (57.77%) followed by DOs (18.95%). Therefore, Keenan concluded that the text frequencies of SU relative clauses with respect to the NPAH "can discriminate syntactically complex authors from syntactically simple ones"(Keenan, 1975, p. 142) The third performance prediction is that authors who are known to use syntactically simple sentences tend to use passive sentences instead of DO relatives. This prediction has also been confirmed by the obtained results. In general, the results that Keenan obtained go with the 'performance predictions' that are drawn from the NPAH. However, the author does not explain the exact parameters that can be used to distinguish simple and complex texts.

One of the NPAH's hypotheses that have been confirmed by Keenan's (1975) study is the SPH. This hypothesis has also been confirmed by Dekeyser (1984) and Romaine (1980) who conducted their studies on the Early Modern English corpus and Scottish English texts respectively. Both of these studies show that the frequency of SU relatives is higher than the frequency of the relative clauses formed on other positions on the NPAH. Apart from the SPH, the results of both studies show that the NPAH's

predictions can be applied diachronically. Dekeyser (1984) and Romaine (1980) found that the use of the *wh*-relative pronoun is in correlation with the NPAH. The authors suggest that the *wh*- relative markers were introduced to the English language in the least accessible positions on the NPAH and expanded to the left positions on the NPAH. On the other hand, deletion of the relative pronouns goes in the opposite way to the WH- expansion, that is from left to right on the NPAH (Dekeyser, 1984). Similarly, Suárez (2012) suggests that the use of the invariable relative marker in early English follows the NPAH. In a study of 'The Helsinki Corpus of English Texts', which represents Old to Middle English, the author found that the distribution of relative clauses with the invariable relativisers *þe* and *þat* is in correlation with the NPAH, although the distribution of relative clauses with *þe* conforms more strongly to the NPAH. The relativiser *þat* is used more than *þe* in Middle English; *þat'*s expansion starts in the least accessible positions of the NPAH. Suárez (2012) does not include all the positions of the NPAH; GEN, OCOMP are not included since the author states that there are no examples found to represent these two positions. IO is also not included, but she does not provide an explanation for the exclusion of IO. Furthermore, she does not explain the parameters upon which the grammatical relations can be distinguished.

Studies conducted by Romaine (1980), Dekeyser (1984) and Suárez (2012) partially support the role of the NPAH in predicting relativisation across different periods of time. A question that needs to be asked is whether applying the NPAH's predictions diachronically can be generalised to other languages such as Arabic. This question can be considered, although this study is not diachronic. This study examines relativisation in varieties of Arabic, and one of these varieties, MSA, is known to be developed from another variety, CA (see section 3.4). Looking at the use of pronoun retention or the use of a relative marker vs. zero relative marker in these Arabic

varieties can provide implications for the diachronic development of these aspects, which can be investigated thoroughly in future research.

Another study that examines the NPAH's predictions in a corpus was conducted by Herrmann (2005), who examines the NPAH's predictions in English dialects. The results she obtained conformed to the NPAH proposed in Keenan and Comrie (1977), but after adjusting it. According to the results the author obtained from the Freiburg English Dialect Corpus by Kortman and Wagner (2005), Herrmann suggests the following hierarchy:

SU> DO> PCOM> GEN> OTHERS

According to Herrmann's hierarchy, the OBL position should be subsumed under PCOM 'the prepositional complement position' which also includes prepositional complements functioning as adverbials, and prepositional complements functioning as *of*-genitive structures. Furthermore, the four positions, which are IO, subject complement, object complement and non-prepositional adverbial, should be put together under the position of 'others', which comes after GEN in the new proposed hierarchy. Moreover, the author believes that OCOMP cannot stand in a separate position and should be included under PCOM. Nonetheless, in Herrmann's study, OCOMP is excluded since there were no instances of this position found in the British dialects. Herrmann's study has shown that the NPAH's predictions can be examined across dialects as well as across languages. However, Hermann's conclusion would have been more convincing if the author had explained whether the results she obtained conformed to the same hierarchy proposed in Keenan and Comrie (1977) and where the deviations, if any, occurred. Since the SU position occupies the first position in Herrmann's hierarchy, Herrmann's study confirms the Subject Primacy hypothesis, as do the other discussed corpus studies (i.e. Dekeyser, 1984; Keenan, 1975; Romaine,

1980; Suárez, 2012). However, other studies suggest that SU relative clauses are not always easier to process and relativise than DO relative clauses. Among these studies is Fox (1987).

Fox challenged the Subject Primacy Hypothesis and presented another hypothesis, which she calls the "Absolutive Hypothesis". Driven by the Preferred Argument Structure Hypothesis of Du Bois (1987), in which he shows that relativisation in discourse conforms to ergativity (i.e. the intransitive subject and direct object are relativised more frequently than transitive subject), Fox distinguishes between intransitive and transitive subjects in her own study. Fox (1987) hypothesizes that if SU relatives are naturally more accessible than other relative clauses, then subjects should be relativised more frequently than other types of relative clauses clause in written and spoken language. Moreover, "if simple texts have the highest ratio of subject to object relatives, then conversation should have the highest ratio of all, being syntactically and structurally simpler than Keenan's written texts" (Fox, 1987, p. 857).

Fox (1987) collected over 100 relative clauses from different kinds of conversations, face-to-face and phone conversations, two- party and multi-party conversations. By analysing her data, Fox observed that SU relatives are not relativised more frequently than other grammatical relations are because the ratio of subject to object is 1:1, 46 relative clauses for each position. The results Fox found show that P-relatives (DO relative clauses) are the most frequent (46 relative clauses) followed by S-relatives (ISU relative clauses; 36 relative clauses), and finally A-relatives (TSU relative clauses; 10 relative clauses). According to Fox:

"[...] the distribution of subject vs. object relatives has more to do with the various functions of each of those kinds of clauses, and with the general treatment of information flow in English, than with cognitive primacy" (Fox, 1987, pp. 857-858).

She explains that the functions that relative clauses have in discourse are either to describe an entity, or to provide an anchor. Most ISU relative clauses are used for the first function which is description because they use stative verbs, mostly the verb 'to be'. On the other hand, the second function, which is anchoring, is carried out by DO relative clauses. The subject in the DO relative clause conveys given information so it helps to link the relative clause to the context. Therefore, the functions of relative clauses in discourse lead to their frequency distribution in the text. Fox conducted her study on spoken discourse only, not written texts. Spoken discourse contains some factors which are not found in written texts, such as body language. Therefore, the results that are obtained in Fox's study cannot be generalized to all kinds of corpora.

The difference between the results obtained by Fox (1987) and those obtained by Keenan (1975) might be attributed to the difference between written corpus and spoken discourse, in which other factors such as the body language contribute to communication. Therefore, there are some studies which have been carried out to validate the NPAH and the AH in relation to corpus analysis. It has been suggested that object relative clauses are more frequent than subject relative clauses in spoken discourse, but not in written texts (Roland et al., 2007, pp. 9-11). Does the frequency of relative clauses in written discourse reflect the NPAH's predictions, whereas the frequency of relative clauses in spoken discourse follows the AH? What are the factors that are found to affect the distribution of relative clauses in written and/or spoken discourse? This section will answer these questions according to the studies which have been conducted so far to investigate the validity of the NPAH and the AH in discourse.

Gordon and Hendrick (2005) claim that the NPAH is reflected in both written and spoken English. The authors collected their data from three different corpora, (Switchboard: Godfrey, Holliman, & McDaniel, 1992; Brown: Kucera & Francis, 1967; CHILDES: MacWhinney, 2000). The results showed that SU relatives (TSU relative clauses+ ISU relative clauses) were produced more often than DO relative clauses in the three corpora. However, by separating the SU relatives into transitive and intransitive, the authors found that the Absolutive Hypothesis was manifested in two corpora, Switchboard and CHILDES. In these two corpora, ISU relative clauses are produced more often than TSU relative clauses. The authors explain that the difference between the percentage of ISU and that of TSU relative clauses was not consistent in these two corpora. They support the NPAH because in general, SU relative clauses are used more frequently than DO relative clauses, which, in turn, are found more frequently than OBL relative clauses. The authors have concluded that the NPAH is not only reflected in both written and spoken discourse, but it is also reflected in adult and child speech, and in face to face as well as telephone conversations (Gordon & Hendrick, 2005, p. 461). However, Gordon and Hendrick's study does not consider the other positions on the NPAH, which are IO, GEN, and OCOMP; in order to support the NPAH, it is necessary that the authors show how relative clauses in all positions (not only in the three positions which they examined) follow the NPAH. Although the SU relatives are found more frequently in all corpora, the results Gordon and Hendrick obtained showed that relativisation in spoken discourse reflects the Absolutive Hypothesis.

Another study was conducted by Hogbin and Song (2007) to investigate the validity of both the NPAH and the AH in written English. The authors argue that the AH is also reflected in written texts. Hogbin and Song collected their data randomly

from English literary texts from two different periods of time, the eighteenth and the twentieth century. The results show that ISU and DO relative clauses are used more than TSU relative clauses. On the other hand, SU relative clauses, TSU and ISU, are found more than other grammatical relations on the NPAH, which is similar to what Gordon and Hendrick (2005) previously observed.

The results obtained by Hogbin and Song reveal the following hierarchy (Hogbin & Song, 2007, p. 229):

S-relatives/DO > OBL > A-relatives > GEN > IO/OCOMP

Thus, in contrast with the other positions on the NPAH, no examples of IO relative clauses have been found in Hogbin and Song (2007), which means this position comes last on the NPAH. This suggests that there is a deviation on the IO position because, according to the NPAH, the IO position should come between the DO and the OBL. In fact, IO relatives have not been detected in previous studies conducted on written English either. In Keenan (1975), this position is included under OBL and Gordon and Hendrick (2005) included only three positions of the NPAH which are SU, DO OBL. The results Hogbin and Song obtained also show a high frequency of OBL relatives, which the authors explain in relation to the descriptive function this type of relative clause performs in the text. Hogbin and Song display the results using descriptive statistics (i.e. percentages of relative clauses); their study would have been enhanced if the authors had used a statistical analysis to examine the significance of the differences among relative clauses in different grammatical positions of the NPAH. Moreover, if different types of texts, for example, fiction and nonfiction texts, had been used instead of fiction only, Hogbin and Song's study would have been more comprehensive.

The distribution of relative clauses in a text is found to be affected by a number of factors. Among these factors are genre, gender and the social class of the writer (Jensen, 1999). Jensen conducted a study on relativisation in two genres of English discourse: diaries and letters. In her study, the author collected data from different regions such as England, Scotland, and Ireland. While geographical regions do not show a significant influence on the distribution of the relative clauses in a text, other factors such as genre do. Jensen points out that the NPAH was represented in diaries more than letters. Gender also appears to be an important factor because Jensen's results show that men's writing adheres to the NPAH whereas women's writing correlates with the AH. The writer's social class is another factor, as the author found that the distribution of relative clauses in the upper class's writing follow the NPAH, in contrast with the distribution of relative clauses in the lower class's writing which conform with the AH. However, the data, in general, support the NPAH because counting both ISU relative clauses and TSU relative clauses under the SU position makes the frequency of SU relatives more than the frequency of relative clauses in other positions of the NPAH.

The effect of genre on the distribution of relative clauses in corpora was emphasized again by Hardy and Milton (1994), who have pointed out that the distribution of relative clauses is determined by the genre in which these relative clauses are used. The authors collected their data from first-person narrative; they also compared their results with other studies conducted on spoken data and data from other fiction and nonfiction texts. They have found that the frequency of ISU relative clauses is highest in informative discourse such as expository writing. On the other hand, DO relative clauses have the highest percentage in exophoric discourse such as conversation whereas TSU relative clauses have the highest percentage in narrative

discourse. The authors suggest that their results are consistent with the NPAH, as both TSU relative clauses and ISU relative clauses make up double the number of DO relative clauses. However, they believe that, in addition to grammatical relations, the functions of the relative clauses and pragmatics of the text are very important factors in favouring certain relative clauses over the others. Genre appears to be an important factor that should be taken into consideration in investigating the frequency of relative clauses in discourse. Therefore, this study will investigate the NPAH and the AH in two genres of Arabic discourse, fiction and nonfiction texts.

This section has presented several studies which have been conducted to re-examine the NPAH's predictions in discourse; the results of these studies are inconsistent. Whereas some of these studies go with the NPAH, others question the validity of the SPH. Additionally, some studies support the NPAH but suggest that a number of factors such as pragmatics, genre, and the writers' gender and social status can cause a particular distribution of relative clauses in discourse. Genre is one of the independent variables in relation to which I examine the NPAH and the AH. The writer's gender is controlled in the choice of the texts that are included in this study in that all authors are male.

2.3.1.2.2 Language Acquisition, Processing and Related Studies

The NPAH has received considerable attention in language acquisition research. In fact, the NPAH's predictions have been examined in first and second language acquisition, yet the results are not consistent, as language acquisition research reveals some support for, and disagreement with, the NPAH. Furthermore, some studies suggest the existence of some factors that might interfere with the NPAH's predictions.

The first and the most important study that relates the NPAH to language acquisition is Keenan and Hawkins (1987). Keenan and Hawkins conducted their study on forty children and forty adults whose L1 was British English. The authors argue that in spite of the differences among the participants regarding age, sex, intelligence, and education, they acquire SU relatives more easily than object relative clauses because the SU position is psychologically more accessible than the other positions on the NPAH. The results obtained by Keenan and Hawkins confirm the SPH, and support the claim that the NPAH represents the psychological order of accessibility.

Most of the studies conducted in the field of second language acquisition (e.g. Croteau, 1995; Doughty, 1991; Eckman et al., 1988; Gas, 1979; Hyltenstam, 1984; Pavesi, 1986) indicate that the order of acquisition of relative clauses by second language learners follows the order of the NPAH. General support with some deviations on certain positions on the NPAH are detected by some of these studies such as Gas (1979), Hyltenstam (1984), and Pavesi (1986). On the other hand, other studies (e.g. Izumi, 2003; Phoocharoensil, 2010) have found that the acquisition of relative clauses by language learners is not predicted by the NPAH only; the NPAH is complemented by other hypotheses such as the Perceptual Difficulty Hypothesis. In general, the focus of all the aforementioned studies is the acquisition of relative clauses in English or European languages.

The SPH was the focus of studies conducted on the acquisition of Asian languages. Some of these studies (e.g. Jeon & Kim, 2007; O'Grady, Lee, & Choo, 2003) have supported the SPH. These studies have found that SU relatives are more easily comprehended and produced by participants than DO relatives. On the other hand, Yip and Matthews (2007) have not supported the claim that SU relatives are always acquired before DO relatives. In their longitudinal study, the authors found that

children acquire and produce DO relative clauses before the SU relative clauses. Thus, even the far-left side of the NPAH, which is represented by the SPH, has not been confirmed by all the studies.

Other studies have considered the factors, other than grammatical relations, that influence language acquisition. Among these factors is animacy of the head (Fernandez-Duque, 2009; Jeon & Kim, 2007; Kanno, 2007; Mak et al., 2002; Ozeki & Shirai, 2007). The authors in these studies claim that animacy of the head noun is an important factor that influences students' performance in the relativisation of SU and DO positions. For example, Kanno (2007) found that the students' good performance on SU relative clauses, as compared to their performance on DO relatives clauses, is conditioned by the existence of semantic clues that help them identifying the grammatical relation of the relativised noun phrase. Among these clues is animacy, as students associate animate with SU and inanimate with DO. The role of animacy in the acquisition of relative clauses is also investigated by Ozeki and Shirai (2007), who found that SU relative clauses referring to animate referents and DO/OBL relative clauses referring to inanimate are easier than SU relative clauses referring to inanimate referents. The important role of animacy has been supported by Hawkins (2007), who mentioned other factors influencing language acquisition such as L1 transfer and semantic factors. Hawkins suggests that the NPAH is

> "..no longer a tenable hypothesis. More interesting questions relate to the role played by parsing, L1 influence, and semantic factors like animacy in the acquisition and use of knowledge about RC structure by L2 speakers" (Hawkins, 2007, p. 349).

Regarding the processing of relative clauses, the processing of SU relative clauses is found to be faster than the processing of DO relative clauses in the Avar

language (Polinsky, Gallo, Graff, & Kravtchenko, 2012). Considering the processing of the two types of SU relative clauses, TSU and ISU, in comparison with DO relative clauses, the authors suggest that there is no difference in processing between TSU and DO relative clauses, whereas ISU relative clauses are processed faster than DO relative clauses. The authors suggest that the processing of relative clauses is motivated by the grammatical functions of relative clauses as well as the case of the relativised noun phrase, which is not relevant to the nominative-accusative languages. The fast processing of SU relative clauses conforms to the SPH of the NPAH.

The grammatical function/position of the head in the main clause is found to have an influential role in the processing of relative clauses. Slobin (1971) suggests that embedded relative clauses represent a kind of interruption which makes the processing of these relative clauses more difficult than other relative clauses. This point is also supported by Kuno (1974), who argued that centre-embedded relative clauses are more difficult to comprehend than right- or left-embedded relative clauses (Kuno, 1974, p. 119); this is found true in pre-nominal as well as post-nominal relative clauses. Both studies imply that the grammatical position of the head of the relative clause in the main clause plays an important role in the processing of relative clauses. Relative clauses that have subject heads[5] are embedded and so they are more difficult to process than relative clauses which have, for example, DO heads. Therefore, according to Slobin and Kuno examples (5) and (6) are more difficult than (7) and (8) which have object heads.

[5] Subject heads refer to heads of relative clauses that occupy the SU position in the main clauses, similar thing with object heads and so on.

5. The man **who plays volleyball very well** is my brother. (SS[6])
6. The man **whom I met** plays volleyball very well. (SO)
7. I met the man **who plays volleyball very well**. (OS)
8. I met the man **whom you like**. (OO)

The data of Sheldon's (1974) study do not support this as the author does not find a significant difference between the performance of the children on embedded and non-embedded relative clauses. Instead, the author suggests that relative clauses where the grammatical function of the head in the main clause is parallel with its grammatical function in the relative clause are easier to process than relative clauses where no such parallelism is found. Therefore, according to Sheldon, examples (5) and (8) are easier to process than (6) and (7). The results that Sheldon obtained suggest that accessibility of relative clauses follow this order: SU-SU, DO-DO, DO-SU, SU-DO. The combined values of SU-SU and DO-SU, which both represent SU relative clauses, are higher than the values of DO-DO and SU-DO, which represent DO relative clauses. Thus, Sheldon's results support the NPAH and the SPH, although the author has not intended to test it. Evidence for the Parallel Function Hypothesis was also found in the Korean language (Clancy, Lee, & Zoh, 1986). In their language acquisition study, the authors found that the children performed better on SU-SU and DO-DO relative clauses than SU-DO and DO-SU, and they considered that to be against the NPAH's predictions. The authors asserted that: "the significant superiority of parallel function SS [SU-SU] and OO [DO-DO] sentences partially contradicts the predictions of the hierarchy [NPAH]" (Clancy et al., 1986, p. 244). The results obtained by Sheldon (1974) do not conform with other studies by Slobin (1971) and Kuno (1974), yet all of these studies

[6] (SS) refers to the grammatical position of the head in the main clause, which is subject, and the grammatical position of the head in the relative clause, which is also subject in this sentence. Similar symbols will be used whenever a reference is made to the grammatical position of the head in the main and relative clauses as is shown in the other examples.

suggest that the grammatical position of the head of the relative clause in the main clause has a big role in the processing, comprehension and production of relative clauses. Therefore, the grammatical position of the head in the main clause is a factor which is worth investigating, and it will be considered in this study.

In conclusion, studies of language acquisition and language processing have shown some deviations on some positions on the NPAH. Indeed, there are some factors which interfere with the NPAH's predictions in the accessibility of the acquisition of relative clauses such as the influence of an L1. Furthermore, some of the factors discussed in the studies above will be considered in this study, such as the parallel function of the head and the position of the relative clause and animacy.

This section has given an account of theoretical and language usage-based studies, which have been conducted to investigate the NPAH in different languages. Studies in this section have shown that the NPAH's predictions have been investigated cross-linguistically and intralinguistically, yet the NPAH has not been supported consistently. Some studies which have been reviewed in this section suggest that there are some factors that might interfere with the NPAH's predictions. These factors might be the reason behind the inconsistent results towards the NPAH. Indeed, it is important to consider these factors in this study.

The next section will review the studies that have been conducted to examine the other hypothesis of the NPAH, which is the use of the relativisation strategies, and whether these strategies are used as predicted by the NPAH.

2.3.2 The Use of Relativisation Strategies on a Continuous Segment of the NPAH

The second hypothesis presented by the NPAH is that relativisation strategies should be used on a continuous segment of the NPAH. If a relativisation strategy is used on the SU position and the OCOMP position, then it should be used continuously on the middle positions too, which are DO, IO, OBL, and GEN. In this section, I will review studies that have been conducted to investigate the use of the relativisation strategies in relation to the NPAH.

Keenan and Comrie (1977) classified relativisation strategies into [+case] strategies and [-case] strategies. The [+case] strategies are case-coding strategies by which the case of the relativised noun phrase is detected. On the other hand, the [-case] strategies do not help in detecting the grammatical relation of the relativised noun phrase. This classification has been criticised by Maxwell (1979), who claimed that classifying relativisation strategies as [+case] and [-case] is a very broad classification. Maxwell suggested another classification which includes *Word Order Strategy,* the *Relative Pronoun Strategy,* and the *Personal Pronoun Strategy*. Although Maxwell's classification is more restrictive, it does not affect the validity of the NPAH. In fact, Maxwell's classification is more difficult to apply cross-linguistically than the original classification proposed by Keenan and Comrie (1977). Findings of other studies (e.g. Cinque, 1981; Song, 1991) go against the NPAH because relativisation strategies in the Italian and Korean languages, respectively, do not conform to the NPAH since these languages do not maintain a continuous segment. However, Song (2003) suggests that relativisation in the Korean language follows the NPAH since the gap strategy, the primary relativisation strategy in Korean, is used for SU, DO and OBL and is also used with other relativisation strategies for OBL and GEN (Song, 2003, p. 140). He states that GEN relative clauses in Korean use the gap strategy as well as the pronoun

retention strategy. Song explains that the gap strategy is used when the 'nominative marked NP', which is the possessed element in the genitive relationship is syntactically and semantically bonded to the predicate inside the relative clause (Song, 2003, p. 147). On the other hand, if the two parts are syntactically and not semantically bonded then the RP is optional. The third case is that the RP must be used if these two elements are neither semantically nor syntactically bonded (Song, 2003, p. 147). He defines bondedness as "the situation in which the NOM-marked NP and the predicate constitutes a close unit" (Song, 2003, p. 146). Song's (2003) study supports the NPAH, yet it also shows how the syntax of the language plays an equally important role in deciding the relativisation strategy used to relativise a certain grammatical relation.

A typological study which is conducted by Wu (2011) on different languages where relative clauses are pronominal shows that relativisation in these languages follows the NPAH. The results the author has obtained show that relativisation strategies follow a continuous segment in pronominal languages except in Chechen which uses the pronoun retention strategy in OBL relative clauses only.

In fact, many studies have focused on the use of the pronoun retention strategy in relation to the NPAH's predictions. Most of these studies are theoretical while a few studies are based on language-usage. The theoretical as well as language-usage based studies will be reviewed in the next section.

2.3.2.1 The Use of the RP and the NPAH

As explained earlier, the pronoun retention strategy is one strategy that is used in relativisation in some languages such as Arabic and Hebrew. When using this strategy, an RP must be inserted in the relative clause to refer back to the head noun

phrase (see section 2.3.2.1). Linguists have attempted to explain the usage of the RP, as they believe that the insertion of the RP in different relative clauses is not done randomly.

Keenan and Comrie (1977) argue that the languages which use the pronoun retention strategy in relativisation apply it to the lower positions in the NPAH such as GEN and OCOMP rather, or more often, than on the upper positions such as SU and DO. Moreover, the use of the RP should take place on a continuous segment on the NPAH. This section reviews some studies that address the use of the pronoun retention strategy in different languages, focusing on the NPAH's prediction.

The correlation between the pronoun retention strategy and the NPAH is specially manifested in Comrie and Kuteva (2005). The authors argue that most of the languages that use the pronoun retention strategy, around 20 languages, use it for the lower positions of the NPAH such as OBL rather than for the upper positions such as SU. However, Babungo is found to be a counterexample to the NPAH because it uses the gap strategy for DO position and the pronoun retention strategy for SU position.

Several other studies show that the use of pronoun retention in different languages conforms to the NPAH (e.g. Bashir, 1982; Cole, 1976; Cole et al., 1977; Lawal, 1987; McCloskey, 1979; Sells, 1987; Suñer, 1998; Willis, 2011). They all confirm that the pronoun retention strategy is preferred to the gap strategy in the positions lower on the NPAH, such as OBL and GEN, in the languages under investigation. Similarly, the data presented by Veenstra and Besten (1994, p. 314) demonstrate that the use of the RP in pidgins and creoles supports the NPAH since RPs are only inserted in OBL and GEN relative clauses, which occupy lower positions on the NPAH.

Additionally, the use of pronoun retention in the German dialects supports the NPAH. This claim is raised by Fleischer (2004), who points out that the pronoun retention strategy is used with the more difficult positions rather than the more accessible ones. The author illustrates that, in spite of the differences among the German dialects, they all use the gap strategy in the SU and the DO positions, while the RP is used for positions lower on the NPAH, such as OBL. The IO position, on the other hand, joins the higher positions in some dialects and the lower positions in others. Thus, although there are differences in relativisation between the German dialects on one side and the standard variety on the other, which has been previously mentioned in Keenan and Comrie (1977), all the German varieties use the pronoun retention strategy according to the NPAH.

On the other hand, the use of the pronoun retention strategy in the English dialects contradicts the NPAH's predictions. In her corpus study, Herrmann (2005) has found that the RP is used eight times in the SU position, but five times in DO relative clauses. The frequency of RPs decreases in lower positions of the NPAH. In contrast with the NPAH's prediction, the use of the RP records the lowest frequency, only one, in the GEN position. Herrmann (2005, p. 73) explains that the NPAH predicts the 'occurrence' of the RP not the frequency of use. That is, the NPAH predicts the grammatical positions in which an RP is required or preferred, but not the number of times it occurs in these positions. The author adds that, in general, RPs are rarely found as they occur only twenty one times, around 4.4 per 100.000 words, which is not significant, and this might contribute to the results that contradict the NPAH's predictions. In fact, Herrmann (2005) has included non-restrictive relative clauses and complex relative clauses, which were not included in Keenan and Comrie (1977). The inclusion of non-restrictive and complex relative clauses might be responsible for the

results that the author has obtained in her study. Both Fleischer (2004) and Herrmann (2005) have demonstrated that the NPAH can be investigated across dialects as well as across languages.

An explanation for the use of the pronoun retention strategy in the least accessible positions has been offered by Hawkins (1999a, 2004), in the framework of his Immediate Constituent Theory of processing. Hawkins suggests that RPs make it easier to process difficult sentences. These complex sentences have long and complex phrases that come between the head and the relativised position. Hawkins (1999a, p. 258) explains that "A pronoun copy makes recognition of the head's role in the relative clause as clear as it can possibly be", so the effort that is needed to infer that role is less than the effort needed in the case of the gap structure. Accordingly, if a gap is used in a certain position on the NPAH, it is predicted to be used for higher positions too. Similarly, if the RP is used for a certain position on the NPAH, it is predicted to be used for lower positions too. Hence, Hawkins presented two predictions regarding the use of RP and the gap strategy.

The NPAH's prediction have not been always the explanation to the use of the RPs in relative clauses. Shlonsky (1992) suggests that the RP is used as a 'last resort' to avoid ungrammatical sentences, where the *wh*-movement is blocked. Other studies (e.g. Aoun, Choueiri, & Hornstein, 2001; Kroch, 1981; Ross, 1967) suggest that RPs are used to avoid violation of island constraints. On the other hand, Heestand, Xiang and Polinsky (2011) suggest that RPs never make English sentences more grammatical. The authors conducted experiments that includes online and offline acceptability judgement tasks. They state that the relative clauses with RPs are judged ungrammatical very quickly, which, they believe, is an indication that the use of RPs in English does not rescue island violation and is not related to processing. They suggest that "the use of

RPs in English is no more than a speaker-centre devise for maintaining coreference" (Heestand et al., 2011, p. 150). Thus, the use of RPs differs from one language into another which suggests that the nature and grammars of a language play an important role in the use of the pronoun retention strategy.

The distance between the head and the relativised noun phrase is confirmed to be an important factor that influences the use of the pronoun retention strategy. Ariel (1990) asserts that the need to insert the RP increases as the distance between the head and the relativised position becomes longer. With reference to Hebrew examples, the author explains that an RP is considered more acceptable in complex sentences. Similarly, Tarallo and Myhill (1983) argue that the distance between the relativised noun phrase and the head has a more crucial role than the grammatical function of the relativised noun phrase in the accessibility of relativisation. Conducting a grammatical judgement test, the authors found that, in languages such as Japanese and Chinese where the relative clauses are prenominal, students accepted ungrammatical RPs in SU relative clauses more than in DO relative clauses. They explain that what motivates the students is that the DO is closer to the head than the SU in these languages. However, the students' performance in relation to other grammatical positions such as IO and OBL does not support the authors' claims. The data presented by Tarallo and Myhill suggest that the rates of acceptance of RPs in IO and OBL positions are close to the rate of acceptance of RPs in the SU position despite the fact that these positions are closer to the head than the SU position. Thus, Tarallo and Myhill's results lead to the conclusion that other factors might affect the accessibility of relative clauses.

The factors that contribute to the choice between the pronoun retention strategy and the gap strategy are discussed in several studies such, as Cole (1976), who suggests that the deletion or retention of the RP in Hebrew can be attributed to the interaction of

different factors such as the case-marking, *Et*-deletion rules[7] and the preposition stranding constraint. Additionally, data collected by Tarallo (1986) support the NPAH since the RPs are found to be used in the lower syntactic positions in Portuguese. However, the author suggests that some factors such as restrictiveness, the distance between the head and gap, and gap embedding, in addition to the syntactic function of the relativised noun phrase, are involved in the choice of relativisation strategy. Similarly, Ariel (1999) suggests that, in addition to the distance between the relativised position and the head, other factors such as the type of head noun phrase, the restrictiveness/ non-restrictiveness of the relative clause and obligatoriness of the relativised noun phrase are found to affect the choice between the pronoun retention or the gap strategy.

Data from language acquisition studies show an increase in the use of the RPs by language learners in the less accessible positions of the NPAH. Among these studies is Labelle (1990), which has shown that the use of the RP increases, with moving down the NPAH, in OBL and GEN. The author also suggests that RPs are used by French children more than adults. Pérez- Leroux and Teresa (1995) have obtained similar results. The authors conducted their study on English as well as Spanish-speaking 3-6-year-old children. The results showed that RPs were used among both English and Spanish children on lower positions on the NPAH such as OBL but not on upper positions such as SU. Other studies (e.g. Gas, 1979; Hyltenstam, 1984; Ioup & Kruse, 1977) show that this strategy is used by students who have and those who do not have this strategy in their first languages, and both types of students use this strategy more for the lower positions of the NPAH. The results of the previous studies clearly agree

[7] *Et* is a preposition in Hebrew which precedes the accusative noun phrases if they are definite so it is considered as a case-marker that distinguishes accusative noun phrases. It is different from the other prepositions in Hebrew as it is restricted to definite noun phrases only.

with the NPAH; however no explanation is offered in these studies for the question of why the use of the pronoun retention strategy by language learners follows the NPAH.

To sum up, the NPAH's predictions about the use of the pronoun retention strategy is confirmed by several studies. However, a number of factors such as the distance between the head and the relativised noun phrase and the type of head noun are also found to affect the use of this strategy. It is important to consider these factors because one of the aims of this study is to investigate the use of RPs in relative clauses in Arabic texts. The use of pronoun retention in Arabic relative clauses has been examined in previous theoretical studies. These studies will be discussed in the following section.

2.3.2.2 Pronoun Retention in Arabic

Keenan and Comrie (1977, pp. 76, 93; 1979, p. 333) claim that the Arabic language uses a [+case] strategy, the pronoun retention strategy, in all positions of the NPAH except in SU relative clauses. Studies on Arabic (e.g.Galal, 2004; Hamdallah & Tushyeh, 1998; Suaieh, 1980) indicate that the RP is not used with SU relative clauses, optional in DO relatives, and obligatory in other relative clauses down the NPAH. Nevertheless, books of Arabic grammar state that RPs can be used in SU relative clauses, and can be deleted under certain conditions. Among these conditions is that the RP can be deleted if it is a subject in the relative clause, and the relative clause is a nominal sentence (i.e. it consists of a subject and a nominal predicate) (Al-muradi, 2001, p. 452).

Previous studies have not provided sufficient explanation for the optional use of RPs in DO relative clauses. Slonsky (1992), who studied Palestinian Arabic and Hebrew, suggests that:

> "there is no free variation between gaps and resumptive pronouns; rather, there is an illusion of free variation that is propped up by the morphological nondistinctness of two crucially different types of relative clauses" (Shlonsky, 1992, p. 448).

Shlonsky built that hypothesis on the assumption that in Hebrew relative clauses[8], where both the RP and the gap are used, the relative marker *še* can be considered as a relative clause complementiser that requires a gap, or a subordinate complementiser that is similar to the Palestinian complementiser , which requires an RP. Since *še* does not have distinct forms for the two cases, the use of RPs vs. gaps appears as optional. On the other hand, Palestinian Arabic relative clauses use only one relative marker, *illi*, so an RP is always required. No optionality is found in DO Palestinian relative clauses.

On the other hand, there is an optionality in standard Arabic, but Shlonsky suggests that what appears as optionality in standard Arabic, "cannot signal a free choice between a resumptive pronoun and a wh-trace [gap]. Rather, the choice seems to lie in the phonetic realization of the resumptive pronoun itself" (1992, p. 458). Thus, Shlonsky's hypothesis denotes that there is always an RP in Arabic DO relative clauses, and the speaker has the choice to produce it explicitly as a pronoun or implicitly (null), and, in the latter choice, the RP can be detected from the rich agreement between the relative marker and the head. However, Shlonsky's hypothesis does not analyse the

[8] The example provided by Slonsky (1992) for Hebrew DO relative clauses is:
ha-ʔiš še ra2iti (ʔoto)
the-man that (I) saw (him)
the man that I saw (p. 452)

basis on which the speaker chooses to show the RP or make the RP explicit, in Arabic relative clauses.

Galal (2004) states that the use and deletion of RPs in DO relative clauses is controlled by semantic and pragmatic factors as well as syntactic factors. Among these factors is the definiteness of the head noun phrase, as the RP cannot be deleted if the head is indefinite. Although, relative clauses in other positions (IO, OBL, GEN, and OCOMP) can also modify indefinite heads, the author does not explain why it is optional only in DO relatives. Suaieh (1980), on the other hand, states that the RP can be deleted in DO relative clauses if the relative clause obeys the rules of island constraints, and if the RP is co-referential with the relative marker and the head. In fact, the RP should be co-referential with the relative marker and the head in all types of relative clauses. Thus, the explanation offered for the conditions that determine the use of RPs in DO relatives is not sufficient. The role of semantic and pragmatic factors has been further supported in other studies such as Jassim (2011, p. 17), who states that RPs are ungrammatical in SU relative clauses that have verbs unless the RP is used for focus. The author explains that "In verbal sentences[9], the subject is expressed by the subject-verb agreement, therefore strong pronouns are optional in such cases and (if produced) they are only used for focus marking" (Jassim, 2011, p. 13).

The same claim has been asserted by Cann, Kaplan & Kempson (2005), who suggest that:

[9] By verbal sentences the author means sentences that include verbs as opposed to equational sentences (for further discussion of verbal and equational sentences see 2.4.2). An example of a verbal sentence is:

Qatal	al-jundi	al-a3da'a.
fought.he	the-fighter	the-enemies.

The fighter fought the enemies.

> "the use of a pronoun in subject position immediately following the relative complementiser can never be justified on the grounds of alleviating cost in processing"(p. 35).

The authors suggest that "strong pronouns" are used as RPs in SU Arabic relative clauses to achieve what they call "additional inferential effects" (Cann et al., 2005, p. 35). The authors explain that the RP is phonologically stressed, as in the following example from Egyptian Arabic (Cann et al., 2005, p. 35):

9.	irra:gil	illi	HUWA	mabsu:t
	the man	REL	he	happy
	The man who he is happy			

On the other hand, in another study on Egyptian Arabic, Eid (1983) suggests that the insertion of pronouns in SU relative clauses has functions other than emphasis. With reference to examples from Egyptian Arabic, the author states that using the pronoun in SU relative clauses imposes a disfavoured reading, which helps to disambiguate the meaning of the sentence, as in example (10) which is used in Eid (1983, p. 289), where the sentence can have two different interpretations, as is shown in the translations. The first interpretation is favoured while the second is less favoured.

10.	ʿali	kallim	il-walad	illi	Ø	(huwa)	šatam-u
	Ali	talked	the-boy	REL	Ø	(he)	insulted-him
	imbaarih						
	yesterday						
	'Ali talked to the boy that insulted him (Ali) yesterday'.						
	'Ali talked to the boy that he (Ali) insulted him (the boy) yesterday'.						

Eid (1983) explains that the verb of the relative clause agrees with the head *il-walad* ‘the-boy’ as well as with the subject of the main clause ‘Ali’. This makes it difficult to decide who is the subject of the verb *šatam* ‘insult’. Thus, inserting *huwa* ‘he’ in the relative clause imposes the second reading where the pronoun is co-referential with the subject of the main clause. However, this explanation contradicts the nature of the RP as being always co-referential with the head of the relative clause.

Other studies such as Badawi et.al. (2004) and Bashir (1982) suggest that RPs can also be omitted in OBL relative clauses. In OBL relative clauses, the absence of the RP should be accompanied with the deletion of the whole prepositional phrase which includes it. Badawi et al. (2004) illustrate that gaps can be used instead of RPs in OBL relative clauses if the head is a time expression. However, the authors do not attempt to explain why it is possible to omit the RP in such cases. On the other hand, following Givón (1979), Bashir (1982) argued that the RP can be omitted together with the prepositional phrase if the pronoun that precedes the head is the same one that precedes the RP (Bashir, 1982, p. 53), as in example (11).

11. ğalas-tu	ʿ**alā**	al-kursīy-i	allaḏī	ğalas-ta
sat-I	**on**	the-chair-GEN	REL(3.SG.MAS)	sat-you
(ʿ**ala**i-hi)				
(on- it)				
I sat on the chair which you sat on.				

The whole prepositional phrase, which includes the preposition *ala* ‘on’ + the RP *hu* ‘it’ can be deleted because the RP is preceded by the same preposition that precedes the head in the main clause *al-kursi-i* ‘the-chair’. On the other hand, if the RP

and the head are preceded by two different prepositions, as in example (12), the relative clause will be ungrammatical if the prepositional phrase is deleted.

12.	ğalas-tu	ʿ**alā**	al-kursīy-i	allaḏī	qafaz-ta
	sat-I	on	the-chair-GEN	REL(3.SG.MAS)	gumped-you
	min-hu				
	from-it				
	I sat on the chair which you jumped from.				

This explanation, which is produced by Givón (1979)[10] and Bashir (1982),

deals with one kind of OBL relative clause in which the function of the relativised noun phrase in both the main sentence and the relative clause is object of the preposition.

In conclusion, previous studies that deal with the use of pronoun retention in Arabic have provided different points of view regarding the types of relative clauses that should or should not use this strategy. Some explanations are based on the pragmatic function of the RP in the relative clause. In the next section, a review of studies that emphasise the pragmatic role of the RP in the interpretation of the relative clause in languages other than Arabic is presented.

[10] Givón suggests that semantics play a more important role than syntax. He provided the following sentences from Hebrew. Although the same preposition is used before the pronoun and before the head, sentence (b) is ungrammatical since it leads to a misleading interpretation:

a.	Hu	yasháv	al	ha-kisé	she-axalti	aláv.
	He	sat	on	the-chair	that-I-ate	on-it.
	He sat on the chair on which I ate.					

b.	Hu	yasháv	al	ha-kisé	she-axalti.
	He	sat	on	the-chair	that-I-ate
	*He sat on the chair which I ate.				

2.3.2.3 The Use of the RP and the Pragmatic Interpretation

Different studies have stated that using an RP in a relative clause gives the sentence an interpretation which is different from that carried out by the same relative clause with a gap. For example, Doron (2011) and Sells (1984, 1987) suggest that using the RP imposes a more specific interpretation whereas the gap suggests a more broad or general interpretation, as in the following Hebrew examples from Doron (2011, p. 305).

13.	Dani	yimca ‘et	ha’isa	se	hu	mexapes	
	Dani	will-find	the woman	that	he	seeks.	
14.	Dani	yimca ‘et	ha’isa	se	hu	mexapes	*ota*
	Dani	will-find	the woman	that	he	seeks	*her*.

Sentence (13) carries two interpretations; the first one is the general meaning, that is there is no particular woman in Dani’s mind, yet he has some characteristics that he likes in a woman, and he is looking for the woman with these characteristics. The second interpretation is that there is a particular woman that Dani is looking for. On the other hand, sentence (14) has only one interpretation; the insertion of the RP *ota* ‘her’ gives a referential meaning, which indicates that there is a particular woman that Dani seeks. This proposal has been supported by Sells (1984, 1987), who suggests that the gap strategy leads to a non-specific reading or what the author calls ‘the concept’, which is not found in relative clauses that use RPs (Sells, 1987, p. 288). However, in both studies, the authors have not considered the role of the verb in imposing a specific or general interpretation, since there are some verbs that suggests a specific interpretation in contrast with other verbs (Diesing, 1992; Enç, 1991).

With reference to Hebrew, Erteschik-Shir (1992) states that one of the functions that the RP has in the sentence is what the author calls "the restrictive focus", as in examples (15) and (16) from Erteschik-Shir (1992, p. 95).

15. hine ha-simla she-kaniti
here the-dress I-bought
Here is the dress that I bought.

16. hine ha-simla she-kaniti ota
here the dress I-bought it
Here is the dress that I bought.

In (15), the addressee knows that the speaker bought a dress. On the other hand, in (16), the addressee knows that the speaker has a number of dresses in mind; therefore using the RP is to focus on the fact that the speaker bought that dress not the other dresses. However, Prince (1990) states that the restrictive function of the RP cannot be found in English and Yiddish. Prince also suggests that the 'concept' proposal cannot be applied to English and Yiddish because of the presence of the definite article which Sells (1987) has not accounted for. In his corpus study of English and Yiddish, Prince has not found supporting examples for the 'concept' and the 'restricting focus' proposals.

The data obtained by Prince (1990) suggest that RPs are mainly used with non-restrictive and indefinite relative clauses. The author explains these results according to what he calls the 'file card' account (Prince, 1990, p. 491). The file card proposal is mainly based on the way the hearer receives and processes the information. In processing sentences, the hearer creates new cards in his mind for indefinite noun phrases, and selects a card from an existing file of cards to represent a definite noun

phrase. In processing non-restrictive and indefinite relative clauses, building new cards or activating existing ones is merely done based on the head of the relative clause. On the other hand, in the processing of definite restrictive relative clauses, the hearer selects the suitable card based on both the head as well as the relative clause. The author states,

> "Resumptive pronouns may occur felicitously just in case the entity evoked by the whole NP is in fact evoked by the head, the relative clause serving simply to predicate some property of that entity, that is, where the appropriate file card has already been independently constructed/ activated" (Prince, 1990, p. 492)

The 'file card' describes how the processing of indefinite and non-restrictive relative clauses are similar, yet it does not explain how this contributes to the use of RPs in these relative clauses rather than with definite restrictive relative clauses.

The studies reviewed in this section attempt to explain the use of RPs on the basis of treating the RPs as ordinary pronouns. Although these studies do not reach a universal explanation of the use of the pronoun retention strategy, they do show that the use of the RP changes the interpretation of the sentence and there is no free variation between relative clauses that use RPs and those which use gaps.

2.4 Summary

A large amount of literature aims to investigate the accessibility to relativisation, in particular the predictions of the NPAH. The NPAH implies that in spite of the differences in relativisation among different languages, all languages follow one hierarchy, the NPAH. According to this hierarchy, the SU relative clauses are the most accessible type of relative clause, which all languages can relativise, while the OCOMP relative clauses are the least accessible. Moreover, the relativisation strategies

that are used in a language, identified by Keenan and Comrie (1977) as [+case] or [–case], should be used in a continuous segment. However, the primary strategy used for SU can be stopped for any lower position on the NPAH.

The NPAH has been investigated in different kinds of studies such as corpus studies and studies of language acquisition, yet there is no general support of the NPAH among these studies. In fact, Fox (1987) is one of the important studies that offer contradictory findings regarding the NPAH. The data Fox obtained from English conversations demonstrate that DO relative clauses and ISU relative clauses are used more frequently than TSU relative clauses. Hence, Fox challenges the SPH proposed by Keenan and Comrie (1977).

Previous studies have suggested that some factors, in addition to the grammatical relations, might affect the accessibility of the relative clauses. Among these factors are the grammatical function of the head of the relative clause in the main clause, animacy and the syntax of the language. Other studies suggest that a number of factors such as, the distance between the relativised position and the head, the type of head noun phrase, and pragmatic factors, might determine the use of the pronoun retention strategy.

To conclude, the NPAH has been re-examined widely through different kinds of studies. However, the research to date has tended to focus on investigating the NPAH's predictions across languages rather than across varieties within the same language. Keenan and Comrie (1977, 1979) have investigated one variety, usually the standard variety, of each language included in their sample. The last point has been critiqued by Fleischer, who points out that including standard German only is "fairly unrepresentative if compared to the overall picture" (2004, p. 236). Thus, a need arises

to investigate the NPAH's predictions in relativisation across varieties rather than across languages. As mentioned previously in this chapter, the NPAH has been applied successfully in two dialectological studies, which are Fleischer (2004) on German and Herrmann (2005) on English. Thus, this study will re-examine the NPAH across three varieties of Arabic, CA, MSA, IA. The next section will show how the investigation of the NPAH in relation to different Arabic varieties is implemented in this study.

2.5 Implications for This Study

Throughout this chapter, it has been observed that several studies have been conducted to examine the NPAH's predictions theoretically and in language use; some of these studies examine both the NPAH and the AH. However, most of these studies focus on relativisation across languages and little attention has been paid to relativisation across varieties. Furthermore, most of the previous studies that examined the NPAH and the AH within languages have focused on one factor at a time, missing other important factors.

This study investigates the NPAH together with AH in light of multiple factors: (1) construction-external factors (variety and genre), and (2) construction-internal factors, which includes characteristics of the head of the relative clause such as definiteness, the grammatical position of the head and animacy. Thus, this study will deal with relativisation in three varieties of Arabic (CA, MSA and IA), and in two different genres: fiction and nonfiction. Relative clauses will be classified according to definiteness into definite and indefinite, and according to animacy of the head into relative clauses that identify human, non-human animate and inanimate, and according into grammatical position of the head in the main clause into different types (see

section 5.4). The gender of the writer is controlled by choosing books written by authors who are all male.

As has been discussed earlier, Keenan (1975) was the first to test the NPAH's predictions using text counts, and since then other corpus studies have used the same methods in testing the NPAH's predictions in a corpus. For example, Fox (1987) has used the same method to test the left end of the NPAH (i.e. subject and object). This study will examine the NPAH's predictions in Arabic texts using methods similar to those used by Keenan and other corpus studies, that is, text counts.

In light of the reviewed literature, this study will attempt to answer the following questions:

1. Do the patterns of relativisation in Arabic follow the NPAH and/or the AH?
2. Is the distribution of relative clauses influenced by the variety in which relative clauses are used?
3. Is the distribution of relative clauses influenced by the genre in which relative clauses are used?
4. Do the characteristics of the heads of the relative clauses such as definiteness, animacy and the grammatical position of the head in the main clause affect the distribution of relative clauses?
5. How do syntactic and semantic factors affect the use of the pronoun retention strategy in Arabic relative clauses?

   The results of this study aim at providing answers to these and other related questions.

Chapter Three: The Arabic Language

3.1 Introduction

The Arabic language is one of the few Semitic languages that survive (Versteegh, 2001). It is the official language of eighteen Arab countries[11]. The number of people who speak Arabic is estimated to be approximately 250 million, and most of these Arabic speakers live in Arab countries. The majority have Arabic as their mother tongue, while others speak Arabic as their second language (Holes, 2004; Versteegh, 2001).

This thesis will mainly focus on relative clauses; however, it is necessary to present the reader with some information about the varieties of Arabic and the structure of the Arabic language for two reasons. First, the Arabic language is characterized by having several varieties that are different from each other on the one hand and different from the Classical/Standard variety on the other so it is necessary to give a general overview about the varieties of Arabic. Second, relative clauses in this thesis are classified according to the grammatical position of the relativised noun phrase, such as SU and DO relative clauses; therefore, it is necessary to provide a background that helps in recognising these grammatical positions in the Arabic language.

This chapter provides a summary of Arabic syntax in 3.2, including a description of verbs, nouns, and pronouns, as well as the structure of Arabic sentences. Section 3.3 will provide an overview of relative clause structures in Arabic, along with the strategies used in relativisation, with particular reference to the pronoun retention strategy. Then, an overview of the varieties of the Arabic language is presented in 3.4,

[11] The countries in which Arabic is spoken are Iraq, Syria, Jordan, Lebanon, Kuwait, Saudi Arabia, Yemen, Oman, Bahrain, United Arab Emirates, Qatar, Egypt, Tunisia, Libya, Morocco, Algeria, Sudan, and Mauritania (Katzner, 2002)

including the relativisation in Arabic dialects. The diglossic situation in Arab countries is presented in 3.5. Finally, section 3.6 will summarise Chapter Three.

3.2 Arabic Grammar[12]

3.2.1 Parts of Speech:

3.2.1.1 The Verb

The verb in Arabic is derived from a lexical root which is defined as " a set of consonants or phonemes in a specific order that embodies a broad lexical meaning" (Ryding, 2005, p. 429). There are two main categories of Arabic verbs: The first category is represented by verbs with roots consisting of three consonants. For example, the verb *kataba*, which means 'to write', has the lexical root *k-t-b.* The second category includes all verbs that are derived from four- consonant lexical roots, for example, the verb *ṣaytar* ' to dominate' has the lexical root *ṣ-y-t-r*; the latter category is less common than the first one (Ryding, 2005).

In general, the verb in Arabic is inflected for gender and number. It distinguishes between masculine and feminine and has three forms for singular, dual and plural. It also distinguishes between the first, the second and the third person (Badawi et al., 2004; Ryding, 2005). The following table shows the inflected forms of the verb *kataba* 'to write' (in the past form):

[12] The grammar of the two Arabic varieties, CA and MSA is similar. Therefore, unless specified, the discussion in the following sections refers to both varieties.

Table 1: Inflection for Person, Number and Gender in the Past Form of the Verb.

Person/Gender	Singular	Dual	Plural
1st MAS.	katab-tu	katab-nā	katab-nā
1st FEM.	katab-tu	katab-nā	katab-nā
2nd MAS.	katab-ta	katab-tumā	katab-tu
2nd FEM.	katab-ti	katab-tumā	katab-tunna
3rd MAS.	katab-Ø	katab-ā	katab-ū
3rd FEM.	katab-at	katab-ā	katab-nna

Table 1 shows that the first person verb is inflected for number but not gender, and there is no distinction between dual and plural. Similarly, there is no gender distinction in the dual verbs of the second and third person.

Verbs in Arabic sentences are inflected for number if they follow the subject, but not if they precede it (Bashir, 1982). This is demonstrated in the following sentences, where the verb is inflected for number only in (18):

17. qatala al-muqātil-ūna bišağāʿah
    fought the-soldier-PRO(3.PL.MAS) bravely
    The soldiers fought bravely.

18. al-muqātil-ūna qātal-**ū** bišağāʿah
    the-soldier- PRO(3.PL.MAS) fought- **PRO(3.PL.MAS)** bravely
    The soldiers fought bravely.

As far as aspect is concerned, there are two kinds of aspect in Arabic: perfective and imperfective (Badawi et al., 2004). However, aspect cannot be separated from tense. Arabic verbs are inflected for present and past tense. The past tense is usually used to indicate completeness of action, or perfective, while the present tense indicates incompleteness of action (Ryding, 2005, p. 440). Hence, there are no specific

structures for continuous and perfect tenses in Arabic. The following table shows the verb *yaktubu* 'to write' (in the present tense) inflected for person, gender and number.

Table 2: Inflection for Person, Number and Gender in the Present Form of the Verb

Person/Gender	Singular	Dual	Plural
1st masculine	a-ktub	na-ktub	na-ktub
1st feminine	a-ktub	na-ktub	na-ktub
2nd masculine	ta-ktub	ta-ktub-āni	ta-ktub-ūna
2nd feminine	ta-ktub-īna	ta-ktub-āni	ta-ktub-nna
3rd masculine	ya-ktub	ya-ktub-āni	ya-ktub-ūna
3rd feminine	ta-ktub	ta-ktub-āni	ya-ktub-nna

As is clear from Table 2, prefixes as well as suffixes are used to indicate the present tense and to indicate person, gender and number. Similar to the past form, the first person verb does not show a gender distinction; nor does it differentiate between dual and plural. The dual forms of second and third person verbs also do not show a gender distinction. In third person verbs, prefixes are used to show gender while suffixes are used to show number.

Arabic verbs are not inflected for future, but future tense can be expressed by adding the prefix *sa-* or the particle *sawfa* to the present form of the verb, as in example (19).

19. **saufa/sa-**'a-ktubu
    FUT- PRES(1.SG)-write
    I will write.

Passivisation in Arabic can be expressed by changing the form of the verb in addition to deleting the subject in the sentence. The verb is known to be passive when a change within the internal vowels has taken place. For example, the past verb k**a**t**a**b is

passivized by changing it into k**uti**b. Changes that take place in the present form of the verb are different from those that take place in the past form. For example, the passive form of the present verb y**a**kt**u**b is y**u**kt**a**b. In the sentences where the verb is passivized, the direct object is considered as a *nāʾib fāʿil* 'substitute of the subject', and it will be in the nominative case instead of accusative, as illustrated in examples (20) and (21).

(Active)

20. ya-ktubu	al-ṭālib-u	al-timrīn-a
PRES(3.SG. MAS)-write	the-student-NOM	the-exercise-ACC
The student writes the exercise.		

(passive)

21. yuktabu	al-timrīn-u
PRES.PASS(3.SG.MAS)-write	the-exercise-NOM
The exercise is written.	

As far as transitivity, verbs are either intransitive, transitive directly or transitive indirectly through a preposition (Badawi et al., 2004; Ryding, 2005). Intransitive verbs do not take direct objects, as in the following example from Ryding (2005, p. 64):

22. ʿāš-au	fī	al-bilādi	al-ʿarabīyah
lived-they	in	the-countries	the-Arab
They lived in Arab countries.			

Transitive verbs can be either monotransitive or ditransitive verbs, as in examples (23) and (24) respectively, which are taken from Badawi et al. (2004). The objects are written in bold.

23. naʿrifu **al-sababa**
we.know **the-reason**
We know the reason. (p. 372)

24. yukalifu al-daulah **mabāliġa** **ḍaḫmah**
It.cost the-state **amount** **huge**
It costs the state huge amounts. (p. 374)

The other type of transitive verbs are "transitive indirectly through prepositions" (Badawi et al., 2004, p. 372) (henceforth this type is called indirect transitive verbs). Indirect transitive verbs are originally intransitive, and they become transitive by the use of the preposition which also changes the meaning of the verb, as is shown in example (25) from Badawi et al (2004).

25. ğāʾa-nī bi-ğihāzin ṣaġīr
came.he-me with-apparatus small
He brought me a small apparatus.

The verb *ğaa'* 'came' is intransitive, yet when it is used with the preposition *–bi* 'with', it is changed to a transitive verb with the meaning of *ğalab* 'brought'.

3.2.1.2 The Noun

Nouns in Arabic can be distinguished from verbs by having either one of the following features:

a. A noun can have the definite article *al* 'the' when it is definite;

b. A noun can have nunation if it is indefinite. Nunation occurs when a final syllable is added to the end of the indefinite noun to indicate its grammatical case. The final syllable is one of the suffixes, *–un*, *-an*, or *–in,* which are used to indicate the following grammatical cases: nominative, accusative, or genitive, respectively[13] (Ryding, 2005).

The first noun of the genitive structure is an exception to the above mentioned two points, since it is indefinite but does not have nunation (Abdelghany, 2010). It instead shows its grammatical case by the usual case markers for the definite nouns (illustrated in table 3). This state will be explained in section 3.2.3.

Nouns in Arabic are inflected for gender, number, and grammatical case. As regards gender, most masculine nouns can be changed into feminine nouns by adding the suffix (*at*), which is attached to the end of the masculine noun. For example, the feminine noun *ṭabīb-**(at)*** 'lady doctor' is formed by attaching the suffix (*at*) to the end of the masculine noun *ṭabīb* 'physician'. However, it is not always the case that feminine nouns are derived from masculine nouns. Indeed, there are feminine nouns that are completely different from their masculine counterparts such as *bint* 'girl' and *walad* 'boy'.

Inflection for grammatical cases in singular nouns is different from that in dual and plural nouns. Singular definite nouns typically use the final vowel (*–u)* to indicate the nominative case, whereas (*-a)* and (*–i)* are used for accusative and genitive,

[13] The following sentences are examples of nominative case in definite and indefinite nouns, respectively:

al-walad-u ḏakīy-un
the-boy-NOM clever-NOM
The boy is clever. Or:
wala-un ḏakīy -un
boy-NOM clever-Nom
A clever boy.

respectively. In singular indefinite nouns word-final nunations are used instead, which are (*–un)* for nominative case, (*-an)* for accusative case and (*-in)* for genitive case. In contrast, most of the dual and plural nouns are inflected for grammatical case by adding suffixes, as illustrated in Table 4.

Table 3: Inflection for Grammatical Cases in Arabic Singular Nouns

Grammatical case	Nominative	Accusative DO, IO	Accusative OBL, GEN
Definite	al-kitāb-u the-book-NOM	al-kitāb-a the-book-ACC	al-kitāb-i the-book-GEN
Indefinite	kitāb-un book-NOM	kitāb-an book-ACC	kitāb-in book-GEN

Table 4: Inflection for Number and Case in Arabic Non-Singular Nouns

Number/ Gender	Nominative	Accusative
Dual: MAS. FEM.	faʾiz -ān winner-DU(NOM) faʾiz -at-ān winner-FEM-DU(NOM)	faʾiz-ain winner-DU(ACC) faiʾz-at-ain winner-FEM-DU(ACC)
Plural: MAS. FEM.	faʾiz -ūn winner-PL(NOM) faʾiz -at-u(n) winner-FEM(NOM)	faʾiz-īn winner-PL(ACC) faʾiz-āt-i(n) winner-PL(ACC)

As can be observed in Tables 3 and 4, the inflection for grammatical cases in dual and plural nouns is different from that in singular nouns. However, the plural feminine nouns are inflected for grammatical case by having the same case markers as singular nouns, which are the final vowels *–u, -a, and -i* with definite nouns, and nunations with indefinite nouns.

3.2.1.3 Personal Pronouns

One of the similarities that MSA and CA share is the personal pronoun system. There are two kinds of personal pronouns in Arabic: free personal pronouns and bound personal pronouns (Badawi et al., 2004; Holes, 2004; Ryding, 2005). Free personal pronouns are pronouns that can be written separately in the sentence and fill the grammatical position held by the subject (Badawi et al., 2004; Holes, 2004; Ryding, 2005), as in example (26):

26. hum al-ṭullāb
    they the-students
    They are the students.

In contrast, bound pronouns are pronouns that cannot stand separately, but must be suffixed to a verb, a noun, or a preposition to fill the grammatical positions held by direct objects, indirect objects, and possessives (Ryding, 2005, p. 298), as in examples (27), (28) and (29), respectively:

27. raʾi-tu-hu qādiman
    saw-I-him coming
    I saw him coming.

28. aʿtai-tu-hu al-kitāb
    gave-I-him the-book
    I gave him the book

29. ḥaqbat-u-hu kabirat-un
    bag-NOM-his big-NOM
    His bag is big.

Most of the free and bound pronouns are inflected for gender, number, and case, as shown in Tables 5 and 6:

Table 5: Inflections in Free Personal Pronouns in Arabic

Person/ Gender	Nominative singular	Accusative singular	Dual Nominative	Dual accusative	Plural nominative	Plural accusative
1st MAS.	ana	īyai	naḥnu	īyānā	naḥnu	īyānā
1st FEM.	ana	īyai	naḥnu	īyānā	naḥnu	īyānā
2nd MAS.	anta	īyaka	antumā	īyākumā	antum	īyakun
2nd FEM.	anti	īyaki	antumā	īyākumā	antunna	īyakun
3rd MAS.	hūa	īyahu	humā	īyāhumā	hum	īyahum
3rd FEM.	hīya	īyaha	huma	īyāhumā	hunna	īyahun

Table 6: Inflections in Bound Personal Pronouns in Arabic

Person/Gender	Singular NOM	Singular ACC/GEN	Dual NOM	Dual ACC/GEN	Plural NOM	Plural ACC/GEN
1st masculine	-tu	-ī	-nā	-nā	-nā	-nā
1st feminine	-tu	-ī	-nā	-nā	-nā	-nā
2nd masculine	-ta	-ka	-tuma	-kumā	-tum	-kum
2nd feminine	-ti	-ki	-tuma	-kumā	-tunna	-kunna
3rd masculine	-a	-hu	-ā	-humā	-ū	-hum
3rd feminine	-at	-hā	-ā	-humā	-ū	-hunna

As can be observed in Tables 5 and 6, singular, dual and plural pronoun(s) of the first person masculine have the same forms as the first person feminine. Moreover, the same dual pronoun is used for the second person masculine and the second person feminine. Similarly, the dual pronouns used for both the third person masculine and the third person feminine have the same form. As for case, all of the pronouns are inflected for nominative and accusative case except (–nā), which is used for nominative and accusative dual and plural.

Sometimes free personal pronouns are used to emphasize the subject (Ryding, 2005, p. 299), as in the example (30):

30. **anā** a-mšī musriʿan
I I-walk quickly
Lit: I I-walk quickly.

As is clear from example (30), the use of the pronoun in the Arabic sentence is for emphasis. The bound pronoun *a-* that is prefixed to the verb *a-mši* indicates that the subject of the verb is the first person singular pronoun, I. Adding the free personal pronoun *ana* 'I' in this case is for emphasis.

3.2.1.4 Prepositions

Prepositions in the Arabic language have different classifications based on semantic and syntactic aspects. Semantically, there are some prepositions that indicate location such as *fī* 'at, in' and *bi-* 'in'; others are used to refer to direction such as *ʾilā* 'to' and *min* 'from'. Some prepositions can refer to both place and time such as *fī* 'at'. Apart from the literal meaning, some prepositions can be used to form a figurative prepositional phrase, such as *ʿalā* in the following examples taken from Ryding (2005, p. 366).

31. ʿalā al-aqal
on the-least
At least

32. ʿalā fikra
on idea
By the way

Syntactically, there are two groups of prepositions: bound and free (Badawi et al., 2004). The bound prepositions are a group of prepositions that cannot stand by themselves in a sentence. Instead, they are prefixed onto nouns, which in turn take the oblique form. These prepositions are *bi-* (in, at, with, by), *li-* (for), *ka-* (like), and *ta-* and *wa-* which both mean 'by'. The preposition in this type consists of one consonant and a vowel, which makes it necessary for these prepositions to be prefixed to the following noun (Ryding, 2005, p. 367).

The other group of prepositions includes free prepositions, those which can stand by themselves in a sentence without any need to be attached to the nouns following them. These prepositions are *min* 'from', *fī* 'in', *ʿalā* 'on, over, above', *maʿa* 'with', *ladā* or *ʿinda* 'at', *ilā* 'to', *ʿan* 'about', *hattā* 'until', and *muḏ/munḏu* 'since'. However, in writing, when these prepositions are followed by pronouns, they should be attached to the pronouns; for example, *fī-hi* (in it or in him), *fī-humā* (in two of them), *fī-hum* (in them), and so forth.

Both of these groups are considered 'true prepositions', and should be distinguished from adverbs of place and time, which are referred to as 'semi prepositions'(Ryding, 2005, p. 366) or prepositionals (Badawi et al., 2004, p. 174). Prepositions, or as they are called in Arabic *ḥurūf al-ǧar*, 'particles of obliqueness', are underived elements, whereas prepositionals are nouns (Badawi et al., 2004). Since prepositionals have some properties of nouns, they may be definite such as, *amam* 'front', and *al-amam* 'the front'. Semantically, the function of prepositionals is to identify the setting in which the action takes place only, whereas prepositions in addition to setting the scene, " can connect verbs with their indirect objects (DOs)" (Badawi et al., 2004, p. 175) in the case of the indirect transitive verb as is explained in 2.4.1.1.

3.2.2 Simple Sentences: Structures and Patterns

In general, there are two types of sentences in the Arabic language: the nominal sentence (*al-ǧumla al-i'smiya*) and the verbal sentence (*al-ǧumla al-fiʿlya*) (Ryding, 2005). A verbal sentence starts with a verb followed by the subject of the sentence, as in the following example:

33. yaktubu al-tilmīḏ-u dars-a-hu
write the-student-NOM homework- ACC-his
The student writes his homework.

On the other hand, the initial element of the nominal sentence is a noun. Nominal sentences consist of two types: equational sentences which do not contain verbs (verbless sentences) (Badawi et al., 2004; Ryding, 2005), and topic comment sentences, which contain a verb, but the verb does not appear sentence-initially.

The equational sentence is the parallel Arabic structure to an English sentence whose main verb is the copula *be*, as in example (34).

34. al-madīnat-u ǧamīlat-un
the-city-NOM beautiful-NOM
The city is beautiful.

The non-expressed verb, *yakūn* 'be', is not used frequently in Arabic sentences (Badawi et al., 2004; Ryding, 2005).

A topic-comment sentence is different from an equational sentence in having a full clause as a comment. The clause is "either an equational or verbal sentence

anaphorically linked to the topic" (Badawi et al., 2004, p. 306), as can be observed in example (35):

35. al-masraḥīya al-šiʿrīya la-hā ḫaṣāʾiṣ tumīyzu-ha
the-drama the-poetic to-it characteristics distinguish-it
poetic drama has characteristics which distinguish it.
(Badawi et al., 2004, p. 328)

Sentence structure in the Arabic language is controlled by rules of concord and agreement (Ryding, 2005, p. 57). Thus, there should be gender and number agreement between the noun and its adjective or its pronoun on one hand and between the noun and its verb on the other. However, there is no agreement required between the verb and the direct object, if there is one, in the sentence. Thus, the subject controls agreement on the verb in the sentence (Bahloul & Harbert, 1993). Agreement between the subject and the verb in a nominal sentence is different from that found in a verbal sentence (Galal, 2004). In the nominal sentence, the verb agrees with the subject in gender as well as number, as in examples (36-39).

36. al-tilmīḏ-u yaktubu wāǧib-a-hu
the-student-NOM he.write homework-ACC-his
The student writes his homework.

37. al-tilmīḏ-āt-u taktubu wāǧib-a-hā
the-student-FEM-NOM she.write homework-ACC-her
The student writes her homework.

38. al-talāmīḏ-u yaktub-ūna wāǧib-āt-a-hum
the-students-NOM write-they(MAS) homework-PL-ACC-their(MAS)
The students write their homework.

39. al-tilmīḏ-āt-u yaktub-nna wāǧib-āt-a-hunn
the-student-NOM write-they(FEM) homework-PL-ACC-their(FEM)
The students write their homework.

In the verbal sentence, on the other hand, the verb agrees with the subject in gender only, as examples (40-42):

40. ya-ktubu al-talāmīḏ-u wāǧib-a-
PRO(3.SG.MAS)-write the-students-NOM homework-ACC-
-hum
-their(MAS)
The students write their homework.

41. ya-ktubu al-tilmīḏ-u wāǧib-a-hu
PRO(3.SG.MAS)-write the-student-NOM homework-ACC-his
The student writes his homework.

42. ta-ktubu al-tilmīḏ-at-u wāǧib-a-
PRO(3.SG.FEM)-write the-student-FEM-NOM homework-ACC-
-hā
-her
The student writes her homework.

Gender and number agreement are not only required in simple sentences, but are also required in complex sentences, such as those including relative clauses, as will be discussed in section 3.2.4.

As can be seen from the previous examples, the subjects of transitive and intransitive verbs, topic-comment sentences and equational sentences all carry nominative case (Ryding, 2005, p. 169). On the other hand, direct objects and indirect objects carry accusative case; obliques and the second part of the annexation, which will be discussed in 3.2.3, carry genitive case (Ryding, 2005, pp. 171-172).

3.2.3 Annexation

Annexation, or what is called *idhafah* in Arabic, is often referred to as the genitive construction, status constructus or construct phrase (Badawi et al., 2004, p. 130; Ryding, 2005, p. 205). Annexation can be defined as follows:

> "two nouns may be linked together in a relationship where the second noun determines the first by identifying it, and thus the two nouns function as one phrase or syntactic unit" (Ryding, 2005, p. 205).

In addition to possession, annexation can be used to convey different meanings or relationships such as identity, partitive, agent, object, compositional or material, measurement, purpose, quotation or title relationship and clause relationship (Badawi et al., 2004, p. 134; Ryding, 2005, pp. 206-210).

Unlike English, in Arabic, annexation or genitive structure does not use a preposition or possessive clitic (Abdelghany, 2010, p. 116), as illustrated in example (43):

43. bāb-u al-maktabat-i
    door-NOM the-library-GEN
    The library door

The first element of the annexation structure which is *bab-**u*** ‘door’, indicates the nominative case by having the final vowel *–u* rather than nunation *–un*, although it is indefinite, and this feature, as discussed earlier in section 2.1.2, exists only in the genitive structure. The second element of the annexation, on the other hand, shows the genitive case by having the final suffix **–i**. Thus, the first element of the annexation structure is always inflected according to its grammatical case, whereas the second element always shows the genitive case marker.

Annexation structure does not always consist of two elements only. Indeed, there are some examples where it involves multiple elements, that is, more than two elements, as in the following example from Abdelghany (2010, p. 118).

44.	ra’i-tu	bait-a	ibn-i	ʿam-i	al-mudīr-i
	saw-1(SG)	house-ACC	son-GEN	uncle-GEN	the-manager-GEN
	I saw the house of the son of the uncle of the manager.				

In the previous sentence, the final three elements are inflected for genitive case (*ibn-**i**, am-**i**, al-mudīr-**i***). On the other hand, the first element *bait-**a*** is in the accusative case because it is a direct object.

3.3 Relative Clauses

The relative clause, in Arabic, is a post-nominal clause that is used to modify an item in a way structurally similar to an attributive adjective. Relative markers are used to introduce relative clauses that modify definite heads only (Bashir, 1982; Ryding, 2005). Hence, when the modified noun is indefinite, no relative marker is used

(Abdelghany, 2010; Suaieh, 1980). The different types of relative clauses are discussed in the following sections.

3.3.1 Types of Relative Clauses

3.3.1.1 Definite-Headed Relative Clauses

The definite-headed relative clause is used to describe and/or identify a definite head. The relative clause is introduced by a relative marker, which usually shows gender and number agreement with the head of the relative clause as is shown in Table 7.

Table 7: The Relative Markers in Arabic

Number	NOM	ACC	NOM	ACC
Singular	allaḏī	allaḏī	allatī	allatī
Dual	allaḏān	allaḏain	allatān	allatain
Plural	allaḏīn	allaḏīn	allawātī or allātī	allawātī or allātī

As can be observed in Table 7, relative markers are inflected for gender and number. A distinction between nominative case and accusative case only appears with the dual relative marker. However, the agreement indicated by the relative markers in Arabic is different from that in English relative clauses as the relative marker in Arabic agrees with the head's grammatical function in the main clause and not with its grammatical function in the relative clause[14]. The following example shows how the relative marker agrees with the grammatical function of the head in the main clause:

[14] There are two clauses in the relative clause sentence: the main clause and the dependent clause, which is the relative clause. For example, the sentence 'the girls I gave the books to are my friends" consists of the two clauses: "the girls are my friends" and "I gave the books to the girls". As a result, the head noun phrase 'the girls' has two functions, it is the SU of the main clause, "the girls are my friends", and at the same time it is the IO of the restrictive clause or dependent clause "I gave the books to the girls".

45. al-walad-ān allaḏān aʿtai-tu-humā
the-boy-PRO(DU.NOM) REL(3.DU.MAS.NOM) gave- I-them(DU)
al-kutub-a aṣdiqāʾ-ī.
the-books-ACC friends-my
The boys whom I gave the books to are my friends

In example (45), the relative marker *allaḏān* agrees with its antecedent, *al-walad-ān* 'the two boys' in number, gender, and case. However, the relativised noun *al-walad-ān* function as IO in the restrictive relative clause. If the relative marker agreed with the relativised noun phrase, that is, IO, it would be in the accusative case *allaḏein.* Therefore, dual relative markers in Arabic show case agreement with the grammatical function of the head rather than with the grammatical function of the relativised noun in the relative clause (Suaieh, 1980). Accordingly, it is suggested that relative markers should be considered as complementiser rather than relative pronouns (Galal, 2004, p. 24; Suaieh, 1980, p. 62). However, the agreement between the relative pronouns and the preceding noun phrases might be due to attraction, where " the relative pronoun is attracted into the case of the noun phrase within the main clause" (Comrie, 1989, p. 154), which is similar to the case in ancient Greek. In this thesis, the term 'relative marker' is used to refer to the relative pronoun/complementiser.

3.3.1.2 Indefinite-Headed Relative Clauses

Indefinite-headed relative clauses are used to modify indefinite heads. In this kind of relative clause, the relative marker must not be used, as is shown in examples (46) and (47):

46. raʾai-tu	rağul-an	Ø	fāza	fī	al-sibāq-i	
saw-I	man-ACC	Ø	won	in	the-race-GEN	
I saw a man who won the race.						

47. *raʾai-tu	rağul-an	allaḏī	fāza	fī	al-sibāq-
saw-I	man-ACC	REL.(3.MAS.SG)	won	in	the-race-
-i					
-GEN					
I saw a man who won the race.					

The sentence in example (47) is ungrammatical since it uses the relative marker *allaḏi* with the indefinite head *rağul-an* ‘a man’.

In fact, the two clauses “ are like two independent sentences implicitly linked because the second refers back to the first” (Ryding, 2005, p. 145) . The link between the two sentences is achieved by the use of the resumptive pronoun instead of the relative marker that is used in headed relative clauses. The resumptive pronoun, which is used in the relative clause is bound to the head (the use of the resumptive pronoun will be discussed further in 2.5.2.1), as in example (48):

48. qaraʾ-tu	qiṣat-an	kataba-**hā**	muʿalim-ī
read-I	story-ACC	wrote-**it**	teacher-my
I read a story that my teacher wrote.			

In this example the pronoun ***–hā*** refers back to *qiṣat-an* ‘a story’, which helps provide a link between the two sentences *qara’tu qiṣatan* ‘I read a story’, and *katabaha mualimi* ‘my teacher wrote’. Traditional Arab grammarians did not consider this type of sentence to be a relative clause since the term ‘relative clause’ was used to refer to

clauses where the relative marker was used. Accordingly, this type was considered as an 'adjectival clause' (Badawi et al., 2004; Suaieh, 1980). Similarly, the relative clause that is used to define a definite noun phrase without using a relative marker is called 'a circumstantial clause'[15] and not a relative clause. Apart from the classical discrimination between adjectival and relative clauses, modern Arab linguists consider both types as relative clauses (Badawi et al., 2004; Ryding, 2005). Furthermore, there are no reasons why traditional Arab linguists distinguish between adjectival clauses and relative clauses because relative clauses are also adjectival. Therefore, both types, definite and indefinite headed relative clauses, will be counted in the data of this study.

3.3.1.3 Free/ Headless Relative Clauses

The headless relative clause is a clause that does not modify an explicitly mentioned noun (Abdelghany, 2010; Suaieh, 1980)[16]. In addition to the relative markers discussed in section 3.3.1.1., this kind of relative clause uses indefinite relative pronouns which are not inflected for gender or number. However, the use of these pronouns is determined by animacy, for example, *man* 'whoever' is used with animate noun phrases, and *ma* 'whatever' is used with inanimate noun phrases. Within the list of the indefinite pronouns, some linguists include *ayu* and *al* which also give a general indefinite meaning (Bashir, 1982, p. 48). Syntactically, relative clauses of this kind can occupy a position in the sentence such as SU or DO similar to any noun phrase

[15] In Arabic, a circumstantial clause is used to refer to "the circumstances under which an action takes place" (Ryding, 2005, p. 283).

[16] Arabic proverbs usually use this kind of relative clause (Suaieh, 1980), as in the following example:

Man jada wajad.
REL works hard.MAS(SG) gets.
Whoever works hard gets good results.

(Abdelghany, 2010, p. 139), as illustrated in the following sentence where the relative clause occupies the DO position:

49. hannaʾa al-mudīr-u allaḏīn nağah-
congratulated the-principal-NOM REL(3.MAS.PL) succeeded-
-ū fī al-imtiḥān-i
-they(MAS) in the-exam-GEN
The principal congratulated those who succeeded in the exam.

3.3.1.4 Restrictive and Non-restrictive Relative Clauses

Relative clauses in Arabic can be restrictive and non-restrictive although the distinction between the two kinds in written Arabic is not as clear as in written English, since there are no punctuation marks (i.e. commas) used with non-restrictive Arabic relative clauses (Al-bazi, 1983, pp. 91-93; Suaieh, 1980, pp. 359-361). Semantically, a restrictive relative clause helps to identify a specific item from a group of items whereas a non-restrictive relative clause provides extra information about the modified item. Accordingly, the distinction between restrictive and non-restrictive relative clauses in Arabic can be made semantically; that is when the referent of a noun is determined based only on the relative clause following it, then this relative clause is restrictive; otherwise, it is non-restrictive.

Having covered the types of relative clauses in Arabic, the next section will discuss the relativisation strategies which are used in Arabic relative clauses.

3.3.2 Relativisation Strategies

There are two relativisation strategies used in Arabic relative clauses: the pronoun retention strategy and the gap strategy[17]. The focus will be on the pronoun retention strategy as it is the preferred strategy in most relative clauses down the NPAH. As is discussed in 1.1.2, the pronoun retention strategy is a [+case] strategy that requires the use of the resumptive pronoun. The following section will explain what the resumptive pronoun is and its use in Arabic relative clauses.

3.3.2.1 The Resumptive Pronoun

The resumptive pronoun (henceforth RP) is a referential pronoun used in the relative clause in place of the relativised noun phrase to refer to the head. It can be used with both definite and indefinite headed relative clauses. Consider the following relative clause, which is an example of a definite headed relative clause:

50.	**al-fatāt**-u	allatī	qābal-tu-***ha***	fī	al-madrasat-
	the-girl-NOM	REL(SG.FEM)	met-I-her	in	the-school-
	-i				
	-GEN				

The girl whom I met in the school.

The resumptive pronoun takes the form of a bound pronoun which is attached to a verb in the case of the object, to a noun in the case of the genitive or to a preposition in the case of the object of the preposition. For subject relative clauses, the RP is a separate pronoun that is inserted in the place of the subject inside the relative clause (Galal, 2004; Hamdallah & Tushyeh, 1998; Suaieh, 1980). It might be argued,

[17] In this thesis, the gap strategy is used to refer to the omission of the resumptive pronoun (RP).

that because the Arabic is a null-language, the RP is carried out by the subject-verb agreement, which is found in the verb of the relative clauses (Aoun, Choueiri, & Benmamoun, 2010). This argument is problematic for two reasons. First, there is not a particular phonological unit in the subject-verb agreement that can be referred to as an RP (Aoun et al., 2010). Gender agreement takes place by adding prefixes and number agreement, if the subject is dual or plural, is achieved by adding suffixes (see Table 2). In this case, it is unclear if the prefixes or the suffixes or both of them can represent the RP. Second, the data of this study, as will be shown in Chapter Seven, reveal some examples where a pronoun is inserted in the relative clauses in addition to the existence of subject-verb agreement, which is obligatory in all Arabic sentences including relative clauses. In considering the agreement as an RP, what is the syntactic function of that separate pronoun? Therefore, in this thesis, I consider the separate pronoun as the RP in subject relative clauses.

Whenever the RP is used, it should agree with both the relative marker and the head in gender and number (refer to Table 5). As is illustrated in example (50), the RP *hu* 'him' agrees with the relative marker as well as the head *Al-ṭalib* 'the-student' in number as they refer to singular, and in gender, as they refer to masculine noun phrases.

The use of the RP makes it possible for the Arabic language to relativise deeply embedded noun phrases that are difficult or impossible to relativise in other languages such as English, as illustrated in example (51) from Suaieh (1980, p. 171).

51. al-ṭālib-u	allaḏī	aʿṭai-tu-***hu***	al-kitāb-a
the-student-NOM	REL(3.SG.MAS)	gave-I-him	the-book-ACC
ḏakīy-un			
clever-NOM			

The student whom I gave the book to is clever/ lit: The male student who I gave him the book is clever.

52. ǧāʾa al-raǧul-u allaḏī yaʿrifu Samir-un
came the-man-NOM REL(3.SG.MAS) know Samir-NOM
al- fatāt-a allatī tazaweǧ-at-hu
the-lady-ACC REL(3.SG.FEM) married-PRO(3.SG.FEM)-him
uḫt-u-hā
sister-NOM-her
*The man that Samir knows the girl whose sister married him came.

Different explanations have been provided in previous studies for the use of RPs vs. gaps. Among the theories that have dealt with the use of the RP are the NPAH by Keenan and Comrie (1977); and the Accessibility Theory by Ariel (1990, 1999). Different factors have been suggested to influence the use of the RP, such as definiteness (Galal, 2004; Suaieh, 1980) and animacy (Erteschik-Shir, 1992) in addition to pragmatics (Doron, 1982; 2004; Joseph, 1980). These theories and claims will be dealt with in detail in chapter seven, where they will be tested against the data of this study.

3.4 Varieties of the Arabic Language

There are a number of varieties of Arabic; these include two standard varieties, CA and MSA, and several spoken varieties. CA is one of the main varieties of the Arabic language. It was the main source of data for Arab linguists who tried to describe and explain the structures used in literary texts, represented mainly by the Qur'an and poetry. It goes back to the sixth century (Ryding, 2005), so, in comparison with other varieties, CA has been studied widely for a long period of time.

The second most important variety of Arabic is MSA, which is used in modern written texts. "The terms 'classical' and 'modern' refer to two diachronic stages in the development of Arabic standard or literary language." (Pashova, 2002, p. 379), yet, there is little agreement on the exact chronological point at which MSA appeared. MSA was associated with writing produced around the middle of the nineteenth century (Holes, 2004). Apart from vocabulary and style, MSA is so similar to CA (Kaye, 2009; Ryding, 2005; Versteegh, 2001) that both of the two varieties, CA and MSA, are referred to in Arabic as *fusha* which means 'the most eloquent one' (Ryding, 2005, p. 4).

There are also a number of spoken varieties[18] or dialects of Arabic such as, Iraqi Arabic, Egyptian Arabic, Saudi Arabic and so on. Indeed, the Arabic dialects differ from CA and MSA at all linguistic levels, including phonology, morphology and syntax (Holes, 2004). Despite the importance of Arabic dialects in communication, few Arab linguists show interest in conducting dialectological studies of Arabic (Versteegh, 2001). However, that is not to say, nor to imply, that there are not any books which discuss the Arabic dialects; rather there are a number of books written on the Arabic dialects by Arab and non-Arab linguists (e.g. Altoma, 1969; Brustad, 2000; Gary & Gamal-Eldin, 1981; Holes, 1990).

Whereas CA and MSA are used mainly in writing, Arabic dialects are usually used in speaking. However, the dialects should not be considered as purely spoken varieties. Instead, Holes (2004) suggests that "it would be more accurate to say that dialectal Arabic is not felt by native speakers to be an appropriate variety of the

[18] A general classification of Arabic dialects is given by Versteegh (2001).There are six main dialects: dialects of the Arabian Peninsula, Mesopotamian dialects, Syro-Lebanese dialects, Egyptian dialects, Maghreb dialects, and African dialects.

language for writing in most texts" (Holes, 2004, pp. 92-93). In fact, dialects are sometimes used in personal letters and informal notes. Thus, dialects can be used in writing as well as speaking in informal contexts, while CA and MSA are used in formal contexts; this emphasizes the diglossic situation in the Arabic- speaking countries, which will be explained section 3.5.

The use of Arabic dialects in writing, particularly in literary texts, has increased recently. Some twentieth-century writers have attempted to use Arabic dialects in their novels and plays especially in the conversations among characters. The use of dialects helps the writer to establish a closer relationship with the reader, and it sometimes helps to differentiate between the educated and non-educated characters (Holes, 2004; Versteegh, 2001). However, the mix between two varieties, MSA and an Arabic dialect, has been very controversial, and many Arabs consider it to be against Arab unity.

The next section will explain how relativisation in Arabic dialects is different from the standard varieties.

3.4.1 Relativisation in Arabic Dialects

The relative clauses in the Arabic dialects are similar to those in MSA in being post-nominal. However, Arabic dialects differ from CA and MSA in terms of the relative markers they use in relativisation. Whereas MSA uses different kinds of relative markers, some of which are inflected for gender and number (see section 3.3.1.1), relative markers in Arabic dialects are not inflected for gender and number (Holes, 2004, p. 284). All varieties use the invariable relative pronoun *illi*, or its variants such as (*halli* or *yalli* for Syrian Arabic or sometimes the short form *ill* in the

Iraqi dialect[19]) for relativisation in all positions (Altoma, 1969; Brustad, 2000; Holes, 1990, 2004).

Most Arabic dialects are similar to MSA in their ability to relativise definite as well as indefinite noun phrases. The relative pronoun *illi* is used with definite noun phrases only (Brustad, 2000), as in examples (37) and (38) which show relativisation in the Iraqi dialect.

53. dag al-bāb al-walad illi ʾīd-***ah*** maksūrah
knocked the-door the-boy REL hand-his broken
The boy whose hand is broken knocked the door.

54. dag al-bāb walad ʾīd-***ah*** maksūra
knocked the-door boy hand-his broken
A boy whose hand is broken knocked the door.

Similar to MSA, the resumptive pronoun *–ah* which is attached to *ʾīd* 'hand', and refers back to the head *walad* 'boy', helps join the two clauses in sentence (54) (see section 3.3.1.2 for an illustration of indefinite headed relative clauses in MSA).

3.5 Diglossia

The word *diglossia* was used by Ferguson to refer to

> " a relatively stable language situation in which, in addition to the primary dialects of the language (which may include standard or regional

[19] There is more than one dialect in Iraq as the spoken dialect in Baghdad is different from the one spoken in the south of Iraq. However, relative clauses in all dialects of Iraq have the same structure. Therefore, no attempt is made in this thesis to distinguish among Iraqi dialects.

> standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety" (Ferguson, 1959, p. 336).

The diglossic situation as explained by Ferguson includes one highly codified variety which is CA (Bassiouney, 2006). Therefore, it is found necessary to expand it to include another highly codified variety that is MSA (Suleiman, 1985) Thus, varieties that are included in the diglossic situation can be classified into levels: (1) high, which are the standard varieties of Arabic, CA and MSA, and (2) low, which is the dialect. Speakers are able to switch between these varieties according to factors such as the formality of the situation.

Arabs tend to use dialectal forms of Arabic in informal conversations with their family and friends, whereas MSA is used in all textbooks, prayers and sermons, lectures, and the media (except in some television soap operas); and CA is used in the Qur'an, religious books and some history books which were written before the nineteenth century. The dialects are the forms of the language that children acquire at home, whereas MSA is the only variety of the Arabic language that is taught formally at schools. In addition to the fact that MSA is considered more prestigious than the dialects, it serves as a means of communication between educated from different parts of the Arab world (Ferguson, 1959; Ryding, 2005).

After Ferguson(1959), different studies have been conducted to investigate the diglossic situation in Arab countries. These studies had been carried out within different frames. Some of them are done within the code-switching frame, where researchers attempted to find out the basis on which people in Arab countries switch from one variety to another, especially from their spoken dialect into the MSA, such as (Bassiouney, 2009b; Eid, 1982, 1988). Other studies investigated the influence of the diglossic situation on Arabic language acquisition, and whether the acquisition of

Arabic language by Arab children is similar o the acquisition of foreign language such as (Eviatar & Ibrahim, 2012; Saiegh-Haddad, 2012). The results of the acquisition studies suggest that diglossia interferes with learning reading and writing in Arabic and the acquisition of standard Arabic is similar to the acquisition of a foreign language.

3.6 Summary

This chapter has given an account of the Arabic language. It has presented a general introduction on parts of speech in the Arabic language and the structure of sentence focusing on the relative clauses. Relativisation in CA and MSA has been discussed in comparison to relativisation in the Iraqi dialect (as an example of an Arabic dialect). Clearly, relative clauses in Arabic dialects are different from relative clauses in CA and MSA. Therefore, relativisation in each and every variety of Arabic is worth studying, and generalisations cannot be made based on examples from one variety.

This chapter has also discussed the diglossic situation in Arabic-speaking countries. It has been explained in this chapter that, in Arabic, there are two main varieties, CA and MSA, which are constant across all Arab countries, and Arabic dialects which are changeable from one country to another.

Chapter Four: Methodology

4.1 Introduction

The main aim of this thesis is to investigate accessibility to relativisation in Arabic by re-examining the NPAH proposed in Keenan and Comrie (1977) in conjunction with the AH offered by Fox (1987). The methods of text selection and data analysis will be described in detail in this chapter. The methods used in selecting the texts are illustrated in section 4.2. Section 4.3 provides a detailed description of the relative clauses that are included in the data collection. An explanation is provided in section 4.4 for the methods of data collection and the statistical methods used for data analysis. The characteristics used to classify relative clauses according to the positions of the NPAH are explained in 4.4.1. Then, the statistical methods which are used in this study will be described in section 4.4.2. Section 4.4.3 clarifies the way by which the statistical analysis is manipulated to describe and analyse the use of pronoun retention in written Arabic. The methods are summarised in section 4.5.

4.2 Selection of Texts

In order to obtain reliable results, all the texts included in this study are classified according to two criteria: variety of the language and genre of the text, as will be discussed in sections 4.2.1 and 4.2.2, respectively.

4.2.1 Varieties of Arabic Included in This Study

Texts have been chosen from three varieties of the Arabic language, CA, MSA and IA, so that the effect that a variety has on the distribution of relative clauses in the texts could be assessed. As has been discussed previously in section 3.4, CA and MSA are both standard varieties of the Arabic language, yet they belong to different chronological periods.

The CA is an older variant of the MSA; the CA texts in this study belong to the period from the tenth to the thirteenth century. Six CA books are included in this study; selection of these texts has been be done by referring to books that discuss Arabic literary texts such as (Allen, 1998; Jayyusi, 2010). In these books, these texts are discussed as classical works. They also provide information about the writers of these classical books such as their birthdates and death dates. The dates of the writers' death, which are inserted at the end of the references in the book list in the beginning of the thesis, further confirm that these works were written in the classical period. The second variety from which the other group of texts is collected is MSA or as it is referred to as the "contemporary variant" of CA (Cuvalay-Haak, 1997, p. 3). Six MSA books are included in this study; these texts are from more recent dates (1996-2008). The third variety is IA. Data on IA is collected from three books; all of these books belong to the twentieth century (1972-1988).

Because the CA is an older variety than the other two varieties, there is a diachronic dimension in the study. However, this study is not designed to test the development of the Arabic language over the time. This study considers CA and MSA as two varieties, as has been done by other linguists such as Ryding (2005), Pashova (2002) and Versteegh (2001). Considering CA and MSA as two varieties is valid

especially with the fact that CA is still in the lives of Muslim Arabs through the recitation of Qur'an.

The factor of variety is concerned with the variety in which the text is written and not the native dialect of the writer. This is applied to two varieties only—CA and MSA, where the variety used by the writer is not considered an important factor as long as the texts these writers have produced represent either CA, MSA. In other words, choosing books that have been written by Iraqi, Egyptian, or other writers who speak other varieties of Arabic is not considered to affect the reliability of the results providing that the texts that are included represent the variety they stand for, which is either A or MSA. On the other hand, the case of choosing IA books is different; books that are used in the sample of Iraqi Arabic are written by Iraqi writers only to ensure that the written conversations represent the Iraqi dialect.

4.2.2 Genres

Texts in this study are classified into two different types according to their genre, fiction and nonfiction. No further classification is attempted. Therefore fiction books are not classified into different kinds (e.g. social, science fiction, romance, etc.), nor are nonfiction books classified into different types (e.g. historic, autobiographic, etc). There are two reasons for using this broad classification in the present study. Firstly, the focus of this study is to give an overview of relativisation in varieties of Arabic in relation to the NPAH and the AH so a general classification into fiction and nonfiction for Arabic texts written in different varieties is enough for the purpose of this study. Secondly, it requires another kind of study to determine a finer distinction between texts. Further classification would be particularly difficult in the case of CA and MSA,

since they stem from different periods, and this would necessitate an investigation of Arabic literature as well as Arabic culture and life at those different times.

The fiction sample is represented by three books which have been selected randomly from each variety. Similarly, the nonfiction sample is represented by three books which have been selected randomly from each variety. A limitation of the aforementioned classification is related mainly to the nonfiction texts. The nonfiction sample includes texts from MSA and CA only, because there are no nonfiction books written in Iraqi Arabic.

Iraqi Arabic is considered as a spoken dialect only, and this is a result of the diglossic nature of the Arabic language (see section 3.5), where there is a "rigid and complementary functional separation of two linguistic codes: the written and the spoken" (Saiegh-Haddad, 2012, p. 45). Thus, there are no texts found written in Iraqi Arabic except the conversations used in novels (see section 3.5). Having the conversations in the Iraqi novels written in the IA is also not common as is confirmed by a note asserted by Al-yasiri (1972, p. 149). In that note, the author states that the committee of the ministry of culture objected to the use of the Iraqi dialect in his novel (Rubāʿīyāt Abu Gāṭiʿ), which they considered against the ministry's regulations of using the standard Arabic in literary works.

4.2.3 Other Issues

The writer's gender has been found to be an influential factor that affects the distribution of relative clauses in texts (Jensen, 1999). In order to eliminate this factor and reduce the effect that gender might have on the distribution of the relative clauses in the texts, the writers selected for this study are male. The texts selected for this study are as shown in Table 8.

Table 8 : Texts Included in the Data Collection for this Study

Variety	Fiction	Nonfiction
CA	1. Alf lailah wa lailah 2. Hayy Ibn Yaqḍān 3. Maqāmāt Al-ḥarīrī	4. Tārīẖ Al-ṭibarī 5. Tārīẖ Ibn Al-aṯīr 6. Tārīẖ Ibn khaldūn
MSA	1. Al-ṭarīq ʾilā tall al-muṭrān 2. Taġrīdat al-bağaʿah 3. Al-manbūḏ	4. Tārīẖ Al-ʿarab wa haḍāratihum fi al-Andalus 5. Al-saif wa al-sīyāsah fi al-Islam. 6. Tārīẖ Al-ʿAarab al-muʿāsir
IA	1. Al-raǧʿ al-baʿīd 2. Al-naẖlah wa al-ǧīrān 3. Rubaʿīyāt Abu Gāṭiʿ	X X X

Reference to any of the above texts is made using the variety abbreviation (e.g. CA) and the number of the book assigned in Table 8. For example, a reference to *Alf lailah wa lailah* is to be made by using the symbol CA, 1, this is followed by the page number such as CA, 1: 95.

Where possible, all relative clauses contained within the selected texts are included in this study. However, this is not the case with novels written in the Iraqi dialect because those novels contain both MSA and IA passages. Only the conversations between the characters are written in IA, so only relative clauses from these parts are counted. In other words, relative clauses that are written in the Iraqi dialect are counted while other relative clauses that are written in MSA are ignored. Similarly, quoted material and poetry, which are sometimes used in the CA and MSA nonfiction books, are not included in the data sample because they could have stemmed from a variety of Arabic different to that in which the book is written.

Some of the selected texts, such as the CA nonfiction books, are extremely long. They run for more than ten volumes, which makes the counting of relative clauses impractical since it is done manually. Therefore, the method used with these books is different from that used with short books, in which all the relative clauses in the book are counted. With long books, data are collected from two hundred pages with an average of one hundred and seventy words per page.

There are two reasons behind using two hundred pages of the long books to collect data from. First, preliminary results showed that the using a lower number of pages did not provide accurate results. The order of relative clauses that is reflected in twenty pages changes dramatically into a different order by the end of fifty pages. Second, using two hundred pages from the long books helps create consistency between the short and long books, because the average length of the short books is two hundred pages. Thus, the relative clauses are counted from two hundred pages that are chosen randomly from the books. If the book consists of ten volumes, then the two hundred pages are divided by the ten volumes. In other words, twenty pages are randomly selected from each volume.

4.3 Relative Clauses

Due to the differences between the traditional definition and the modern definition of the Arabic relative clause (refer to 3. 3.1.2), the semantic definition that is used in Keenan and Comrie (1977) will be used as a basis in this study because that definition enabled the researchers to identify relative clauses successfully in around fifty languages. Therefore, it will be used in this study to determine the kind of relative clause that should be included in the sample. This definition is as follows:

> "We consider any syntactic object to be an RC if it specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified, called the domain of relativization, and then restricted to some subset of which a certain sentence, the restricting sentence, is true. The domain of relativization is expressed in surface structure by the head NP, and the restricting sentence by the restricting clause, which may look more or less like a surface sentence depending on the language". (Keenan & Comrie, 1977, pp. 63-64).

To illustrate, in the relative clause '*the girl that I met in the library'*, the head noun phrase '*the girl*' specifies the domain of relativisation, a set of all girls. The restricting clause is '*(that) I met in the library*'.

It is clear from the above definition that the data of this study will include restrictive relative clauses only. Non-restrictive relative clauses will be excluded for two reasons. First, using Keenan and Comrie's definition to identify relative clauses suggests including restrictive relative clauses only. Second, previous studies such as Tarallo (1986) and Ariel (1999) indicate that restrictiveness is one of the factors that interfere with the NPAH's predictions, so it is important to include similar relative clauses (i.e. either restrictive or non-restrictive relative clause) in order to control this variable and to make a reliable comparison with the results of previous studies.

Additionally, Keenan and Comrie's definition suggests that only clauses are included in the data collection rather than adjectives and nominals. Therefore, structures that identify persons or things using adjectives, active participles, and passive participles will be excluded from data collection, in spite of their describing and restricting function. Moreover, structures that do not contain either relative markers or verbs are not included in the sample used in this study although they might be interpreted as reduced relative clauses. For example, the sentence, 'The girl sitting on

the chair, might be interpreted as ‘the girl who is sitting on the chair’. Although Keenan and Comrie (1977) included these structures in their data sample, considering them as relative clauses is still controversial among linguists. Therefore, the structures that are included in this study should have at least one of the following characteristics: first, a relative clause should contain a relative marker; second, a relative clause should contain a verb.

The data include both definite relative clauses (those identifying definite heads) and indefinite relative clauses (those identifying indefinite heads). Keenan and Comrie (1977) did not include indefinite relative clauses. However, the authors did not provide an explanation for excluding this kind of relative clause from their data. On the other hand, other studies such as Fox (1987) and Gordon and Hendrick (2005) did include indefinite relative clauses because the authors did not find any apparent reason why indefinite relative clauses should be excluded. As far as Arabic relativisation is concerned, traditional Arab grammarians consider clauses that identify indefinite heads as adjectival and not as relative clauses (see section 2.6.1.2). However, I will include indefinite relative clauses in this study. In fact, definite relative clauses and indefinite relative clauses have the same function, which is identifying a thing or a person. Consider the following examples:

55. tamma al-qabḍu ʿalā al-muğrimi allaḏī irtakaba
done the-arrest on the-criminal REL(3.SG.MAS) committed
ğarīmata qatli aḫ-ī-k
crime murder brother-GEN-your
The arrest of the criminal who murdered your brother is done.
(MSA, 3: 51)

In (55), the relative clause *allaḏī irtakaba ǧarīmata qatli aḫīk* (who murdered your brother) helps to identify *al-muǧrim* (the criminal) and to distinguish him from other criminals. Similarly, relative clauses with indefinite heads have the same function as in the following example:

56.	wa	yufāḫiru	bi-anna-hu	ṣāra	mufakir-an	taḫšā-hu
	and	boast	in-that-he	became	thinker-ACC	fear-him
	al-daula					
	the-government					

And boast himself that he became a thinker the government fears.
(MSA, 2: 83)

Accordingly, the data collected for this study consists of all restrictive relative clauses that are found in the selected texts. It might be argued that Arabic does not have punctuation marks, such as commas in English, to distinguish non-restrictive relative clauses from restrictive ones. That is certainly true; nonetheless, Keenan and Comrie's (1977) semantic definition is helpful in determining whether a relative clause is restrictive or non-restrictive. Therefore, if the relative clause in the text helps the reader to identify a thing or a person from a set of things or persons, then it is considered restrictive. On the other hand, if the reader can identify the thing or person the writer talks about without the need to use the relative clause, then the relative clause is non-restrictive.

Occasionally, it is difficult to decide whether some relative clauses are restrictive or not, despite the semantic definition, as in example 57.

57. tārik-īn ḫiām-a-hum wa ġanāʾim-a-hum
leaving-PL tents-ACC-their and spoils-ACC-their
allatī lam yatamakan-ū min ḥamli-
REL(3.SG. FEM) did not able-PL from carrying-
- hā maʿa-hum
-them with-them
Leaving their tents and spoils that they could not carry with them.
Leaving their tents and spoils, which they could not carry with them.
(MSA, 4: 59)

In (57), it is not clear whether they carried some of the tents and spoils and left the ones they could not carry, or whether they left all the tents and spoils and the clause here is to clarify more about the reasons behind leaving them (i.e. because they could not carry them). The assumption of the first meaning leads to a restrictive interpretation, while assuming the second meaning leads to a non-restrictive interpretation. Such relative clauses are excluded from this study. This exclusion is unlikely to have a significant effect on the results of this study because they are estimated to be less than 3% of the data sample.

Headless relative clauses are excluded from analysis for two reasons. First, they do not have the functions of restricting and describing in the same way headed relative clauses do because they do not have head nouns to describe. Secondly, most similar studies (e.g. Fox, 1987; Hogbin & Song, 2007; Keenan & Comrie, 1977) have not included headless relative clauses due to the syntactic difference between the two types. Therefore, this study examines headed-restrictive relative clauses only.

As a result, the basic definition of the relative clause included in this study is that it is a simple restricting clause that contains an explicitly expressed verb and is used to identify a definite or indefinite head.

4.4 Test Methods

Counting relative clauses is the principal method used in this study. Relative clauses in the sample texts are counted and then classified according to the positions of the NPAH. Then the percentage of relative clauses formed on each position is worked out depending on the number of relative clauses found in the texts. This method is chosen because it has been proven to be effective in previous studies, including the two major ones (Fox, 1987; Keenan, 1975) where the frequency of relative clauses in each position is implemented as a measure of the accessibility of that position to relativisation.

Section 4.4.1 discusses the classification of relative clauses into different kinds according to the NPAH. Then, in 4.4.2, the statistical methodology used to analyse the data is explained. Finally, section 4.4.3 focuses on the use of the pronoun retention strategy in different positions and how it can be analysed statistically.

4.4.1 Grammatical Positions of the NPAH

It is necessary to decide the grammatical positions of the hierarchy that are included in the study before any count takes place. The original hierarchy proposed in Keenan and Comrie (1977) is as follows:

1. SU > DO > IO > OBL > GEN > OCOMP

However, Keenan (1987) joins two positions, IO and OBL, under the OBL position because he claims that the two positions are similar. In this study, on the other hand, there is an intention to test the NPAH as originally proposed in Keenan and Comrie (1977) because these two positions are different in the Arabic language (see

sections 4.4.1.3 and 4.4.1.4). Therefore, relative clauses in IO position and OBL position are counted separately.

According to Keenan and Comrie (1977), the SU position is the most accessible position followed by the DO and the other positions going down the hierarchy, IO, OBL, GEN, OCOMP. Fox (1987), on the other hand, claims that intransitive subject and direct object are more accessible to relativisation than the transitive subject (refer to section 3.3). Therefore, the assumption of SU relatives as the most accessible position cannot be taken for granted especially in light of other studies that agree with Fox's claims (e.g. Gordon & Hendrick, 2005; Hogbin & Song, 2007; Roland et al., 2007). Thus, in this study, in order to test both the SPH and the AH, SU relative clauses will be further classified into transitive subject relative clauses (TSU) and intransitive subject relative clauses (ISU). Accordingly, hierarchy (1) will be tested as follows:

2. ISU + TSU > DO > IO > OBL > GEN > OCOMP

According to the AH presented by Fox (1987), the predicted hierarchy is:

3. ISU + DO > TSU > IO > OBL > GEN > OCOMP

In this study, both hierarchies (2) and (3) will be considered and the data of this study will show which of these hierarchies is reflected in Arabic relative clauses and in which variety and which type of text.

4.4.1.1 Transitive vs. Intransitive Subject Relatives

Traditionally, there are two morphological tests that are used to identify subjects: (1) case and (2) triggering agreement in the sentence (Bickel, 2010, p. 402)[20]. These tests have continued to be used until the present (Witzlack-Makarevich, 2010, p. 2) . Accordingly, in Arabic an NP can be a subject if it, first, has the nominative case (Ryding, 2005, p. 169), and second, if it triggers number, person and gender agreement in the verb (Bahloul & Harbert, 1993), although subject verb agreement in the nominal sentence is different from the verbal sentence (see section 3.2.2). Examples of subject-verb agreement are in the invented examples (58) and (59).

58. yaktub u al-tilmīḏ-u al-qiṣat-a
    he.write the-student-NOM the-story-ACC
    The student writes the story.

59. taktubu al-tilmīḏ-at-u al-qiṣat-a
    she.write the-student-FEM-NOM the-story-ACC
    The student writes the story.

As is clear from examples (58) and (59), the verb is changed from *yaktubu* 'he.write' into *taktubu* 'she.write' since the subject in (59) is feminine. On the other hand, the DO remains the same.

As is implied in the previous section, SU relatives will be classified into TSU and ISU. TSU includes relative clauses whose verbs take one or more objects as in examples (60) and (61), respectively:

[20] There are several theories that have been founded to identify the grammatical relations across languages using parameters other than those used in this study (Witzlack-Makarevich, 2010). It is beyond the scope of this study to discuss these theories. The morphological tests are enough to identify the grammatical relations, in particular the SU and the DO, for the purposes of this study. Further tests will be discussed and used as needed throughout this chapter.

60. la nurīdu ḥilm-an yukabilu-nā
not we.want dream-ACC tie up-us
We do not want a dream that ties us up.
(MSA, 2: 249)

61. al-sabab allaḏī ḫaʿala al-madīnat-a tartʿibu
the-reason REL[3.SG.MAS] made the-city-ACC frighten
min-i
from-me
The reason that made the city frightened of me. (MSA, 1: 62)

The ISU relative clause, on the other hand, includes two types of relative clause. The first type is a relative clause whose main verb is intransitive, as in example (62).

62. šifi-t al-youm waḥid chān yištuġul wiy-yay bi-l-bank
saw-I the-day one was work with-me in-the- bank
I saw today one who was working with me in the bank.
(IA, 1: 80)

The other type of ISU relative clause is derived from equational sentences (refer to section 3.2.2). The equational sentence is a verbless sentence that consists of a subject and a predicate which "may take the form of different classes of words and phrases: nouns, predicate adjectives, pronouns, or prepositional phrases" (Ryding, 2005, p. 59). It is similar to English sentences that employ the verb 'be' as a main verb, as in example 63, where the clause between brackets is considered a relative clause since it is preceded by a relative marker:

63. arġab an yatimma irğāʾa baiʿi al-bilād
I.want to happen postpone selling the-land
allatī [fi al-wādi]
REL(3.SG.FEM) [in the-valley]
I want to postpone selling the land which is in the valley.
(MSA, 3: 70)

Alternatively, some sentences do not have relative markers since they have indefinite heads. It is difficult to assess whether these sentences are relative clauses given that they have neither verbs nor relative markers. They could instead fit into another category such as the modifying noun phrases. An example of this kind of sentence is (64).

64. flam ya.qbala-hu min-hum aḥad siwa rağul-un min
so-not he.accept-him of-them one except man-NOM of
banī Ziyād
Family Ziyad
No one accept him except a man who is a member of Zyad's family.
(CA4: 2149).

Although the previous sentence is translated into a relative clause in English, it is not a relative clause in Arabic because *min banī Zyād* 'of Zyad's family', which defines the head noun *rağul-un* 'a man', is a prepositional phrase and is not preceded by a relative marker.

It might be argued that the previous prepositional phrase can be considered to be a reduced relative clause based on the assumption that originally this sentence had a relative marker but the relative marker was deleted because the head was indefinite

(refer to section 3.2.4.1.2). That is, if the head were definite, then a relative marker would be inserted before the prepositional phrase, which makes the prepositional phrase a relative clause. However, this is just an assumption and there is no way to confirm it. Moreover, neither traditional nor contemporary Arab linguists consider similar examples to be relative clauses (Badawi et al., 2004). Hence, in an attempt to obtain consistent results, and due to the uncertainty about the classification of these sentences, such sentences will be excluded from analysis in this study.

Another important point that needs to be illustrated in distinguishing between TSU and ISU relative clauses is that of transitive and intransitive verbs. It is not always easy to distinguish between transitive and intransitive verbs. There are some Arabic verbs that are followed by prepositional phrases (PP) instead of a DO, yet they are considered as indirect transitive verbs (see section 3.2.1.1). These verbs are "semantically transitive but govern their object indirectly" (Danks, 2011, p. 105). In such cases, the prepositional phrase seems to be so connected to the verb that the sentence will not be semantically and syntactically complete using the verb alone (Qabawa, 1989, p. 273). It is argued that the prepositional phrases that are used after this kind of verb take the place of the DO (Qabawa, 1989).

The distinction between the two kinds of intransitive verb is illustrated by Hawkins (1999), who classifies prepositional verbs into dependent and independent using different factors. Among these factors is what he calls 'lexical dependency', which means "that either V or PP requires access to the other for the addition of semantic and/or syntactic properties that are not assignable to each on its own" (Hawkins, 1999b, p. 251). Hawkins further explains:

> "If some PP is obligatorily required by an intransitive verb or if there is a dependency between V and P in either direction, we might view

such properties as sufficient for classification of the relevant PP as a complement" (Hawkins, 2004, p. 113).

By way of illustration, Hawkins (1999b) differentiates between two intransitive verbs, 'wait for' and 'count on' as used in the sentences 65 and 66:

65. The man waited for his son in the early morning.

66. The man counted on his son in his old age" (Hawkins, 1999b, p. 242).

The following two tests are used to distinguish between the verbs in the previous sentences:

a. Verb entailment test:

"If [X V PP PP] entails [X V], then assign Vi. If not, assign Vd, indicating (by a "v" subscript) which PP(s) is/ are required in addition to V for an entailment to go through (e.g., [XVd PPv PP] entails [X Vd PPv])" (Hawkins, 1999b, p. 242).

b. Pro-verb entailment test:

"If [X V PP] entails [X Pro-V PP] or [*something* Pro-V PP] for any pro-verb sentence listed below, then assign Pi, if not, assign Pd. Pro-verb sentences: X *did something* PP; X*was* PP; *something happened* PP; *something was the case* PP; *something was done (by X)* PP" (Hawkins, 1999b, pp. 242-243).

Hawkins (2004, pp. 113- 114) explains that the meaning of the sentence '*I counted on John*' does not entail '*I counted*', whereas the meaning of the sentence '*I waited for John*' entails '*I waited*'. In fact, the preposition 'on' should be attached to the verb 'counted' to make the sentence mean 'depend on'.

As for the pro-verb entailment test, the two verbs are different, since the verb in the sentence '*I waited for John*' can be replaced by '*I did something for John*', whereas

verb in the sentence '*I counted on John*' does not mean, and cannot be replaced by, '*I did something on John*'.

The application of the previous tests to Arabic sentences shows that the prepositional phrases that are considered DOs by Arab grammarians are the same ones which Hawkins (1999b, 2004) considers as dependent. This can be illustrated in example (67) where the verb used is a transitive verb although it is followed by a preposition:

67. inna-hu masʾūl kabīr fī al-šarikah allatī
that-he official big in the-company REL(3.SG.FEM)
taqūm bi-l-ḥafur
FEM.stand for-the-digging
He is a senior officer in the company which performs the digging.
(MSA, 3:179)

Applying the verb entailment test to example (67) shows that the verb *taqoum* 'stand' in this sentence is transitive because it depends on the prepositional phrase to imply the meaning of the sentence. The intention is not to say '*the company stands*', but instead to say '*the company performs*'. This cannot be achieved without the use of the prepositional phrase because the verb *taqūm* alone means 'stand', while *taqūm bi-...* means 'perform'. The sentence *al-šarikah taqūm bi-l-ḥafur*' 'the company performs the digging' does not entail *al-šarikah taqūm* 'the company stands'.

Similarly, the pro-verb entailment test further proves that the verb *taqūm* in (67) is transitive. The sentence: *al-šarikah taqūm bi-l-ḥafur* 'the company performs the digging' indicates that the company does something and not 'the company does something + pp' *bi-l-ḥafur*, 'does something in the digging'. Therefore, this verb and

similar ones are considered transitive verbs in this study although they are believed to be less transitive than verbs which are followed directly by a direct object (Danks, 2011, p. 113).

For the purpose of this study in order to distinguish between the TSU and ISU, the previous tests will be applied first. Then the verb will be checked in Al-Milani (1979), where the author has collected the Arabic verbs that have become transitive through the use of prepositions from the most reliable Arabic dictionaries (e.g. Al-fayyuumi, 1987; Al-raazi, 1988; Al-zamakhshari, 1984).

4.4.1.2 The DO Position

Following the morphological tests of case and verb agreement, an NP can be considered a DO in Arabic if it has the accusative case, and it does not trigger agreement in the verb. That is, the verb agrees with the SU of the sentence in number, person and gender but not with the DO, as has already been shown in sections 2.4.2 and 4.4.1.2. In this study, I have included both DO noun phrases which follow the transitive verb directly, as in example (68), and DO noun phrases which do not follow the verb directly, instead they are attached to a preposition which immediately follows the transitive verb, as in example (69).

68. rafaḍ	siğāra-tī	allatī	qaddam-tu-hā
refused	cigarette-my	REL(3.SG.FEM)	presented-I-it
ʾilai-hi			
to-him			

He refused my cigarette that I have presented to him.
[MSA, 2: 10]

In sentence (68), the noun phrase *siğara-tī* ‘my cigarette’ is directly affected by the verb *qaddam* ‘presented’ and it directly follows the verb. This is clear from the position of the RP, which is suffixed to the verb *qaddam-tu-hā* ‘presented-I-**it**’.

69. ini ʾaraa hāḏā al-raǧul allaḏī ǧiʾt-um
I see this the-man REL(3.SG.MAS) came-you(2.PL)
bi-hi laisa maʿa-hu ṯāni
in-him has no with-him second
I see that man you brought has no one with him. [CA, 4: 119]

It is clear that the verb in this sentence is transitive through the use of the prepositional phrase since the original sentence is *ǧi’t-um bi-l-raǧul* ‘came-you in-the man’. Applying the tests discussed in section 4.4.1.1, I have found that if the verb *ǧi’t-um* is used without the prepositional phrase *bi-hi*, it means ‘you came’, but it does not imply that you brought the man with you. This indicates that the prepositional phrase is a complement rather than adjunct since the relation between the verb and the prepositional phrase is so strong that the sentence will be syntactically and semantically incomplete without it.

Al-raǧul ‘the-man’ in sentence (69) appears to be an object of the preposition since it has genitive case and follows a preposition. However, Arab grammarians consider the preposition in sentence (69) and similar prepositions as particles of transitivity by which intransitive verbs become transitive (Al-anbari, 1957, p. 86; Al-zajaji, 1979, p. 109; Badawi et al., 2004, pp. 372, 380-383; Sibawayh, 1988, p. 38). Through these prepositions, the effect of the verb is transferred onto their objects. Thus, a*l-raǧul* ‘the-man’ in sentence (69) is considered an object rather than an object of the

preposition despite the fact that it has genitive case and it follows a preposition. Consequently, the relative clause in (69) is considered a DO relative clause, and so are other relative clauses if the verbs are proven to be transitive.

4.4.1.3 The IO Position

Identifying the IO position in this study is made mainly on a semantic basis because, in Arabic, both DOs and IOs, if they are not preceded by prepositions, carry the accusative case; therefore the morphological tests are not always helpful in distinguishing between the DO and IO. Before identifying the IO relative clauses, it is helpful to define the ditransitive structure that will be included in this study.

> A ditransitive construction is defined here as a construction consisting of a (ditransitive) verb, an agent argument (A), a recipient-like argument (R), and a theme argument (T). [...] The most typical ditransitive constructions contain a verb of physical transfer such as 'give', 'lend', 'hand', 'sell', 'return', describing a scene in which an agent participant causes an object to pass into the possession of an animate receiver (=recipient) (Malchukov, Haspelmath, & Comrie, 2010, p. 2)

In this study, an NP is an IO if it is has the role of recipient or receiver in the ditransitive structure. IO relative clauses include those sentences in which the head noun phrase occupies the position of IO, or, as it is called in Arabic 'the second object'[21] inside the relative clause. The IO can be found in two types of structures: 'prepositional dative' and 'ditransitive dative' (Bashir, 1982, pp. 104-105; Ryding, 2011).

[21] In Arabic, the IO is called the 'second object', which may includes arguments other than IO because ditransitive verbs in Arabic also include causative types (to make x do y) and transformational type (to make x into y' and 'to consider x to be y) (Badawi et al., 2004, p. 374). These types are not included in this study.

In the ‘prepositional dative’ structure, the IO follows the preposition *li* ‘to’ or ‘for’ and carries the genitive case (Ryding, 2011, p. 290), as in (70), where the sentence before relativising the IO is a*l-ğundī sallam al-risālā li-lqāʾid* ‘the soldier delivered the letter to the leader’.

70.	al- qāʾid	allaḏī	sallam	la-hu	al-ğundī
	the-leader	REL(3.SG.MAS)	delivered	to-him	the-soldier
	al-risālah				
	the-letter				

The leader to whom the soldier delivered the letter to.

In the second structure, the ‘ditransitive dative’, the IO, which carries the accusative case, lacks the preposition and follows the verb immediately (Ryding, 2011, pp. 290-291), as in (71).

71.	al- qāʾid	allaḏī	sallama-hu	al-ğundī	al-
	the-leader	REL(3.SG.MAS)	delivered- him	the-soldier	the-
	risālah				
	letter				

The leader to whom the soldier delivered the letter.

In example (71), the sentence before relativising the IO is a*l-ğundī sallam al-qāʾid al- risālah* ‘the soldier gave the leader the letter’.

4.4.1.4 The OBL Position

The OBL relative clause refers to relative clauses in which the relativised noun phrase is an object of the preposition. That is to say, the head of the relative clause is co-referential with the object of the preposition in the relative clause. In this study, all

relative clauses with prepositional phrases are included under the OBL category except prepositional phrases which are proven to be complements not adjuncts. The distinction between adjuncts and complements depends strongly on classifying verbs into transitive and intransitive, which has been discussed in 4.4.1.1. If the entailment tests prove that a prepositional verb is transitive, the prepositional phrase will be considered a DO, despite the fact that it is apparently an object of the preposition. Similarly, despite the similarity between some of the IO relative clauses, where the IO carries the genitive case, and the OBL, which also carries the genitive case, IO will be distinguished from OBL by referring to the definition of the ditransitive structure (see section 4.4.1.3). An example of an OBL relative clause is (72).

OBL relative clauses are classified into different categories according to the category of the head of the relative clause and the semantic function the preposition performs in the sentence. These categories are place, time, means, condition, possession, source, manner, cause, accompaniment, result, concern, and basis (refer to Appendix A for examples of these categories).

72. fa-ẖarağ-tu anā min al-makān allaḏī kun-
then-went out-I I from the-place REL(3.SG.MAS) was-
-tu fī-hi sirran
-I in-it secretly
Then I went out secretly from the place in which I was.
[MSA, 1: 301]

4.4.1.5 The GEN Position

Genitive relative clauses can be recognised when the second part of the annexation structure, which carries the genitive case, is relativised (refer to section

3.2.3 for more information about Arabic genitive structure). The two parts of the annexation structure can both be nouns, as in sentence (73) which is an example of a GEN relative where the noun phrase *al-manzil* (‘the house’) in *sālat-u al-manzil* (‘the lounge of the house’) is relativised.

73. tawaqaf-tu	amām	al-manzil	allaḏī		aśʿala
stopped-I	in front of	the-house	REL(3.SG.MAS)		turned on
-t Ilene	al-nūr	fī	sālat- i-hi		
FEM Ilene	the-light	in	lounge-GEN-its		

I stopped in front of the house whose lounge Ilene turned on the light in.
[MSA, 1: 246]

On the other hand, there are other examples where the first element of the annexation structure is a verbal noun which is similar to a participle or an infinitive in English, as in example (74).

74. ġeir	ābihin	bi-afruʿi	al-šağari	allatī	yatawālā
not	care	in-branches	the-trees	REL(3.SG.FEM)	successively
ṣuqūṭi-hā					
falling-their					

He does not care about the tree’s branches whose falling is successive.
(MSA, 2: 14)

The example in (74), is considered as a GEN relative clause because verbal nouns behave like nouns in their ability to be annexed to their agents or objects (Badawi et al., 2004, p. 237). Therefore, GEN relative clauses in the two cases, where the two parts of the annexation structure are nouns, a verbal noun and a noun, are both included in this study.

As far as the semantic relation conveyed by the GEN relative clauses, Keenan and Comrie (1977) included only possessors under the GEN position. However, in this study, the GEN position covers a number of relationships besides possession (Ryding, 2005, pp. 206-210). The classification of GEN relative clauses into different kinds of relationships follows the classifications found in previous studies (Al-dahdah, 2007; Janet & Watson, 1993; Ryding, 2005). Among these relationships are Patient, Identity, Agent, Adverbial and Partitive (see Appendix A for examples of these types).

4.4.1.6 The OCOMP Position

Arabic is one of the few languages that allow relativisation of object of comparison noun phrases (Bashir, 1982; Hamdallah & Tushyeh, 1998). This position includes relativisation of the noun phrase after the preposition *min* 'than', which in turn follows a comparative argument and carries the genitive case, as is shown in example (75).

75. qad nazala bi-nā qawm lam nara
already came down in-us people not we.see
qauman qaṭ aḥsana min-hum
people never better than-them
People who we have never seen people better than came to our house.
[CA. 4: 103]

Since OCOMP uses one of the prepositions which is used with OBL, a careful distinction should be made between the two positions. Bashir (1982) suggests one important point that helps in the distinction between the OBL and OCOMP relative clauses. In the OBL position, the position of the prepositional phrase is flexible; that is

the prepositional phrase is able to move to different places in the relative clause (before the verb, after the verb and at the end of the sentence), as is shown in examples (76-78), where the prepositional phrases are in bold type. In contrast, the prepositional phrase in the OCOMP is always at the end of the relative clause as shown in example (75) above.

76. al-muʿalim-u allaḏī istaʿar-tu al-kitāb-a
the-teacher-NOM REL(3.SG.MAS) borrowed-I the-book-ACC
min- hu
from-him
The teacher whom I borrowed the book from.

77. al-muʿalim-u allaḏī istaʿar-tu **min-hu**
the-teacher-NOM REL(3.SG.MAS) borrowed-I **from-him**
al-kitāb-a.
the-book-ACC
The teacher whom I borrowed the book from.

78. al-muʿalim-u allaḏī **min-hu** istaʿar-tu al-
the-teacher-NOM REL(3.SG.MAS) **from-him** borrowed-I the-
kitab-a
book-ACC
The teacher whom I borrowed the book from.

4.4.2 Statistical Analysis of the Data

After finishing the data collection, I analyse the data statistically using multi-level Poisson regression analysis. This method of data analysis has proven to be a good way of analysing textual frequencies (Baayen, 2008), and it is capable of considering multiple factors at the same time. Multi-level Poisson regression analysis has also been used in similar studies such as Keune, Ernestus, Van Hout and Baayen (2005), where

the authors examined the frequencies of certain words used in different registers across different regions. Keune et al.'s (2005) study is similar to the present study, since the latter looks at the frequency of relative clauses used in different types of texts across different varieties.

The same method is used to work out the differences among the frequency of relative clauses formed on each position on the NPAH and the adjacent ones. The multi-level Poisson regression models that are used in this study help in controlling any differences among the books which might take place due to the writers' different styles. Considering books as a random factor reduces the chance that the results obtained in this study are attributed to individual stylistic differences. The analysis of the data also takes into consideration the construction-external and construction-internal factors that are included in this study. As far as construction-external factors, which are variety and genre, using multi-level Poisson regression analysis helps to determine whether the differences among varieties and between genres are significant, and whether varieties and genres play a significant role in the distribution of relative clauses in Arabic texts.

As far as the construction-internal factors, the first factor that is considered in this study is definiteness. Definiteness has been examined in a number of previous studies such as Fox (1987), Fox and Thompson (1990), Gordon and Hendrick (2005) and Hogbin and Song (2007) among others. These studies have focused on grammatical relations and the tendency to use definite noun phrases in certain grammatical positions rather than the others. This study, however, deals with definiteness in relation to different variables such as the grammatical position of the head in the relative clause (i.e. the NPAH), the grammatical position of the head in the main clause, animacy, and construction-external factors. The relative clauses are classified into definite and

indefinite (see section 4.3). Then, the sum of definite noun phrases in each position of the NPAH is worked out, and using multi-level Poisson regression analysis, I analyse whether the difference between definite and indefinite relative clauses is significant in relation to the grammatical positions of the NPAH and the other factors included in this study.

Another construction-internal factor that is considered in this study is the grammatical position of the head in the main clause. Previous studies (e.g. Kuno, 1974; Sheldon, 1974; Slobin, 1971) suggested that the grammatical position that the relativised noun phrase occupies in the main sentence has an important role in the processing and acquisition of relative clauses. This claim is tested statistically by classifying the relative clauses according to the function of the relativised noun phrase in the main clause and in the relative clause, which is similar to the classification used in Fox and Thompson (1990). This classification results in having different groups with different patterns such as SU-SU or DO-SU. Then, I use multi-level Poisson regression analysis to find out whether there is any significantly high tendency for heads of relative clauses towards occupying certain grammatical positions rather than others in the main clause. I also test whether there is a positive relation between the grammatical position of the head in the main clause and other factors that are tested in this study such as variety, genre, definiteness and animacy.

The third factor that is considered in this study is the animacy of the head of the relative clauses. A number of studies (Fox & Thompson, 1990; e.g. Hawkins, 2007; Jeon & Kim, 2007; Kanno, 2007; Mak et al., 2002; Ozeki & Shirai, 2007; Roland et al., 2007) have suggested that the animacy of the head of the relative clause plays an important role in the processing, distribution and acquisition of relative clauses. In this study, relative clauses are classified according to the animacy of their heads into

human, non-human animate and inanimate. The results of animacy are analysed statistically using multi-level Poisson regression analysis to find out whether the differences among the three categories are significant, and also to decide whether there is a positive relation between animacy of the head on one hand and other construction-internal and construction-external factors on the other.

4.4.3 The Pronoun Retention Strategy

As discussed in sections 1.1.2, 2.5.2.1 and 3.3.2.2, Arabic is one of the languages that use the pronoun retention strategy in relativisation, especially on the lower positions on the NPAH. Keenan and Comrie (1979, p. 333) claim that "pronouns are obligatorily retained in the NPrel [positions relativised] except for subjects". However, other studies on the Arabic language assert that RPs are not used in some positions other than SU, such as DO (Galal, 2004) and OBL (Badawi et al., 2004; Bashir, 1982)(refer to section 3.3.2.2). The following sentences are examples of DO and OBL with an RP in (79 and 81) and without in (80 and 82), respectively, the RPs appear in bold:

79. mā hīya al-muqaūwimāt allatī
what PRO(3.SG.FEM) the-element.PL REL(3.SG.MAS)
kān yamlika- **hā** Muʿāwia ḥattā yusalima-hu al-
was He.has-**them** Maʿawia to he.handdown-him the-
Sham?
Sham?
What were the elements Maʿawia had that he (the ẖalifa) handed down the Sham to him?
[MSA,5: 97]

80. wa	aśʿal-ū	fī	kuli	nāḥia	min	al-ḥaṭab	
and	lit-they	in	each	side	of	the-wood	
allaḏī		ğamaʿ-ū-Ø[22]			la-hu		
REL(3.SG.MAS)		collect-they-Ø			for- him		

and lit the fire in every side of the wood which they collected. [CA, 4: 84]

81. ma	abraka	al-sāʿa	allatī	ʾatait-a	fī-	
what	blessed	the-hour	REL(3.SG.MAS)	came-you	in-	
-ha						
-it						

What a blessed hour in which you came. (CA.1:177)

82. maḏā	waqtun	ṭawīl	lam	ušāhidu-hu	Ø-Ø
went	time	long	not	I.saw-him	Ø-Ø

It has been a long time in which I have not seen him. (MSA, 3:132)

In this study, a distinction is made between relative clauses that have RPs and those that do not have RPs. Then the sum of the RPs in each position on the NPAH is worked out, and the percentages are calculated. The results of the pronoun retention strategy are analysed using descriptive analysis (i.e. using percentages only) and not the multi-regression analysis because the numbers of relative clauses that do not use the pronoun retention strategy are very small in some positions, not allowing for regression analysis. Once the percentage of the RPs is calculated, the validity of the NPAH's constraints will be tested. First, according to the percentage of RPs, this study determines whether the use of the pronoun retention strategy follows the NPAH's predictions without any violation; the NPAH predicts the pronoun retention strategy should be used on the lower positions of the NPAH, OCOMP, GEN and OBL, more than on upper positions, SU, DO and IO. Second, this study determines whether the use

[22] Ø indicates that this would have been the place of the RP if it had been used. If Ø is used twice, as in example (82), this indicates that there is another argument that is not used in the sentence; in the case of the relative clause in (82), it is the preposition.

of this strategy applies to a continuous segment of the NPAH. This study also tests the previous claims about the use of the pronoun retention strategy. Among these claims are the Accessibility Theory, the distance between the head and the relativised noun phrase, definiteness, and animacy. To determine the effects of genre and variety on the use of the RP in relative clauses, a comparison is made between the two genres and among the three varieties.

4.5 Summary

This chapter has presented the methodology which is used to investigate the predictions of the NPAH proposed by Keenan and Comrie (1977), together with the Absolutive Hypothesis presented by Fox (1987), in different varieties of the Arabic language. First, the selection of texts has been discussed. Then, this chapter has examined the types of relative clauses that are included in this study and the way in which they are counted and statistically analysed. The frequency of different types of relative clauses are calculated and multi-level Poisson regression analysis is used to analyse the effects of construction-external factors, genre and variety, in addition to construction-internal factors, definiteness, the grammatical position of the head in the main clause, and animacy, on the distribution of relative clauses in Arabic texts. Finally, this chapter has shown how the use of the pronoun retention strategy is analysed in light of the NPAH and other claims from previous studies.

Chapter Five: Results and Analysis

5.1 Introduction

As explained in the previous chapters, Keenan (1975) claims that the distribution of relative clauses in written texts follows the NPAH, which represents accessibility to relativisation cross-linguistically (Keenan & Comrie, 1977). On the other hand, Fox (1987) suggests that the functions which relative clauses have in a text determine how frequently those relative clauses are used, and so changes the far left side of the NPAH to conform with the AH. This study aims, first, to examine the distribution of relative clauses in Arabic texts in relation to these two hypotheses; second, to examine the influence of construction-external and construction-internal factors, discussed in the previous chapters, on the distribution of relative clauses in Arabic texts.

This chapter will present the results of the corpus study. It will show how strongly the predictions of the NPAH and/or the AH are borne out by the patterns of Arabic relativisation. This chapter is organised as follows: Section 5.2 will present the general results and counts. Then the results will be displayed first in relation to the NPAH in section 5.2.1 and after that in relation to the AH in section 5.2.2. The effects of definiteness, the position of the head of relative clause in the main clause and animacy on the distribution of relative clauses are presented in sections 5.3, 5.4 and 5.5, respectively. Then, a comparison among the results of the three varieties is presented in section 5.6 and a comparison between the two genres in section 5.7. Finally, a summary of the results chapter is presented in section 5.8.

5.2 Assessment of the NPAH and the AH

The Arabic texts included in this study yielded 2785 relative clauses. These relative clauses appear in the descending order: ISU (789), DO (749), TSU (601), OBL (475), GEN (161), OCOMP (8), IO (2). Figure 2 shows the preference for ISU relative clauses over other types of relative clauses.

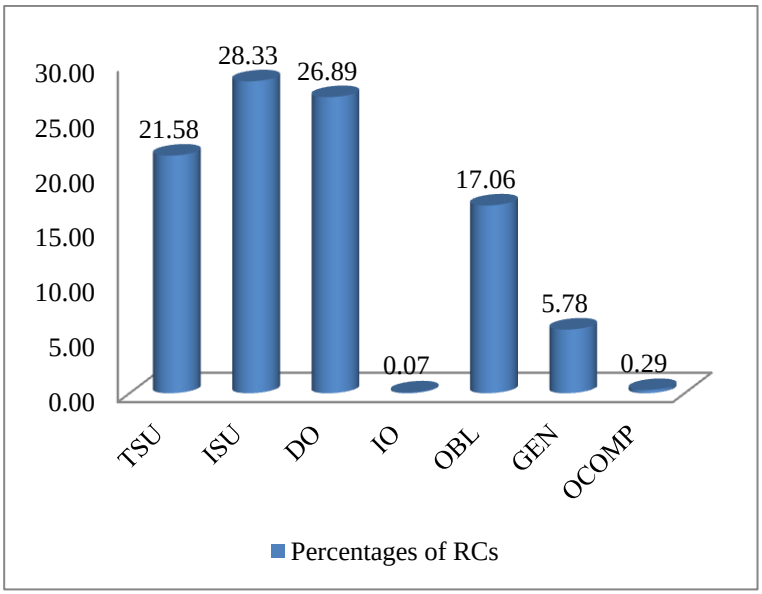


Figure 2: Distribution of Arabic relative clauses in written texts

This study uses multi-level Poisson regression analysis for statistical analysis of the relative clause counts, as such analysis can take account of both categorical and continuous variables at the same time. The models that are created in this study control the differences among texts by considering the text as a random factor. This controls for variation due to text-specific factors. Because of the small numbers of IO relative clauses (2) and OCOMP (8), as will be explained in section 5.2.1, these two positions are excluded from this and other statistical models in this study, as is shown in Table 9.

Table 9: Model 1: Mixed Effect Poisson Regression for the Distribution of Relative Clauses in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.266	0.133	9.500	<0.001 *
TSU	-0.362	0.185	-1.953	0.050 .
DO	-0.036	0.131	-0.279	0.780
OBL	-0.547	0.127	-4.300	< 0.001 *
GEN	-1.573	0.164	-9.577	< 0.001 *
Def	0.987	0.043	23.105	< 0.001 *
INANIM	1.059	0.043	24.379	< 0.001 *
NONFIC	0.542	0.127	4.281	< 0.001 *
IA	-1.273	0.223	-5.698	< 0.001 *
MSA	-0.279	0.127	-2.206	0.027 *[23]

Model 1 tests five predictors: (1) relative clause types with the SU position split into ISU and TSU, (2) definiteness, (3) animacy, (4) variety and (5) genre; these predictors appear in the first column of the table. The dependent variable is the count of relative clauses of the relevant category. Rows 2-5 of Table 9 show a comparison between ISU and other relative clauses on the NPAH. Levels of each predictor are coded by alphabetical order; for example, in the case of the models in this study, all relative clauses on the grammatical positions of the NPAH would be compared to the DO relative clauses since DO comes first in alphabetical order. In this and the following models, alphabetic characters (a, b, c, etc.) are joined to the names of the grammatical positions for ordering purposes. For example, ISU becomes a.ISU, and TSU becomes b.TSU and so on down the NPAH to make the results appear in the order of the positions in the NPAH. Therefore, Model 1 compares relative clauses in all grammatical positions to ISU.

[23] The asterisk in the tables that show the results of the multi-regression models refers to the fact that the value is significant. Therefore, it is different from the asterisk used with some sentences in the previous chapters, which indicates that the sentence is ungrammatical.

The second column in Table 9 shows the estimate value for each type of relative clause; all relative clause types have negative estimates, which indicates that the model predicts other relative clause types to have smaller counts than ISU. However, this result will be changed when interactions with animacy, variety and genre are considered in sections 5.5, 5.6 and 5.7, respectively. These interactions could not be included in Model 1 because the data collected in this study is not enough to test all the possible interactions in one model.

A *p*-value of less than 0.05 is considered as significant, represented by an asterisk in the table. The dot indicates that the value is approaching significance. The *p*-values in the last column of Table 9 indicate that the differences between ISU relative clauses and relative clauses in other grammatical positions except in the DO position are significant, as is shown in Figure 3. The difference between ISU and DO relative clauses is not significant (*Wald's z*= -0.279, *p*=0.7804); this indicates that the difference between ISU and DO being due to chance cannot be ruled out. Therefore, there is no clear evidence that ISU and DO are different. The last five rows represent the results of definiteness, animacy, variety and genre, which will be explained in sections 5.3 and 5.5, 5.6 and 5.7 respectively. The reason why these variables are included in this model is to test the distribution of relative clauses with the presence of other fixed effects so that the relationship between the frequency of relative clauses and other variables is controlled.

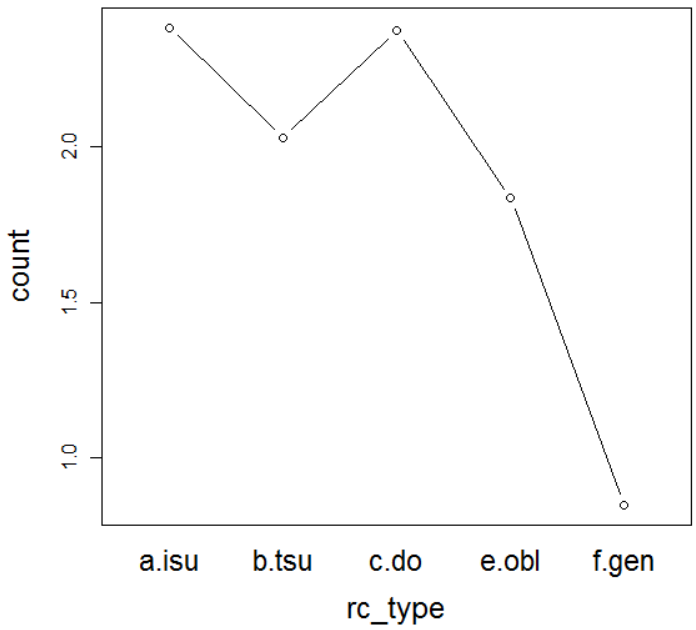


Figure 3: The distribution of relative clauses in Arabic texts

Figure 3 has five points that show the grammatical positions tested in Model 1, ISU, TSU, DO, OBL, GEN. The count axis shows how frequently the relative clauses are used on different positions within the regression model. As is shown in Figure 3, the DO position is close to the ISU position, which is the most frequently used. On the other hand, the DO position is higher than the TSU position with a significant difference (*Wald's z*= 3.117, *p<0.001*), and there is no statistical difference between TSU and OBL (*Wald's z*= 1.128, *p*=0.259), the difference between the OBL and GEN is significant (*Wald's z*= -7.562, *p*<0.001) (see Tables 1 and 2 in Appendix B).

5.2.1 The NPAH

The data of this study show that the combined values for the TSU and ISU, which makes up (49.91%), are significantly higher than the values of other positions of the NPAH (see Table 3 in Appendix B), as is shown in Figure 4.

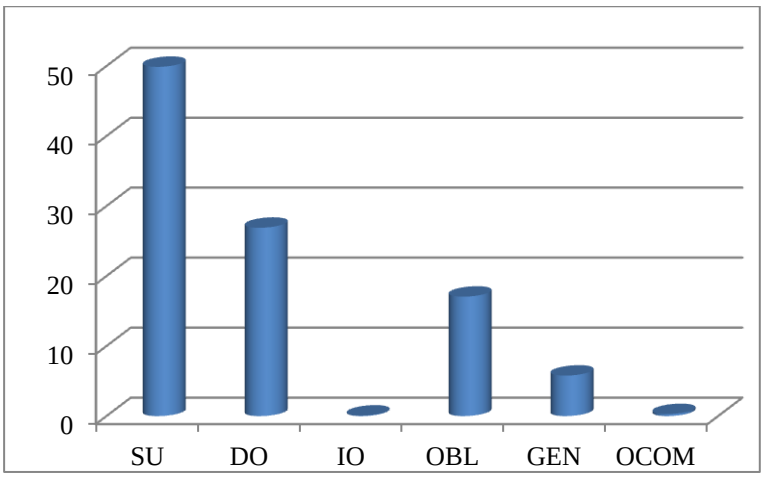


Figure 4: The distribution of relative clauses in Arabic texts when TSU and ISU are combined into SU

Considering the other positions on the NPAH, the DO position comes second with the percentage (26.91%), which also supports the NPAH. However, a deviation can be noticed in the position of IO since only two IO relative clauses have been found. The percentage of the IO position is (0.07%); this causes the IO position to come last in the order of the relative clauses found in the Arabic texts. Despite the fact that the IO position is the third position of the NPAH, it is outnumbered by the lower positions on the hierarchy, OBL (17.03%), GEN (5.78%) and OCOMP (0.28%). There are eight OCOMP relative clauses found in the texts; this makes the percentage of OCOMP relative clauses slightly higher than the percentage of IO, yet both of them have a percentage of less than 1%. In relation to the NPAH's predictions, the infrequent

occurrence of the OCOMP relative clauses is as expected since this position comes at the end of the NPAH, yet it is not expected for the IO position. A discussion of these results in relation to the NPAH's predictions will be provided in section 6.3.

The OBL position comes third after the SU and DO and the difference between the OBL position and the SU position is significant (*Wald's z*=-10.336, *p*<0.001) (see Table 3 in Appendix B). Regarding the classification of OBL relative clauses (refer to section 4.4.1.4), results show that categories of OBL relative clauses are found in the following order from the highest percentage: place 185 (39.02%), time 106 (22.36%), mean 75 (15.82%), condition 61 (12.86%), possession 21 (4.43%), source 12 (2.53%), manner 4 (0.84%), cause 3 (0.63%), accompaniment 2 (0.42%) , result 2 (0.42%), concern 2 (0.42%), and basis 1 (0.21%), as shown in Figure 5 (for examples of these categories refer to Appendix A).

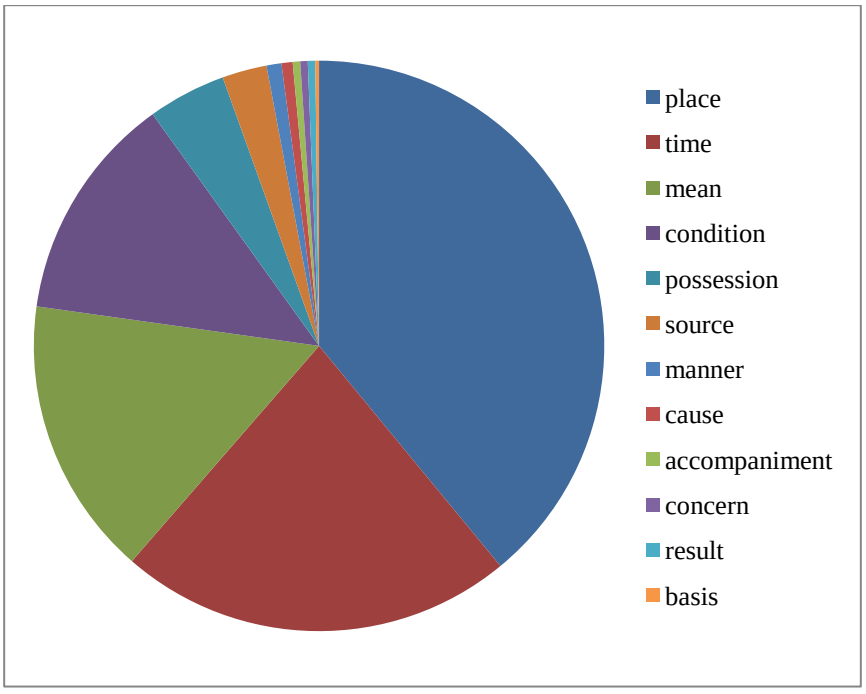


Figure 5: The distribution of OBL relative clauses according to their types

The GEN position comes fourth after the SU, DO and OBL positions, and the difference between the GEN position and the SU position is also significant (*Wald's* z=-14.921, p< 0.001) (see Table 3 in Appendix B). As far as the distribution of different categories of GEN relative clauses (refer to section 4.4.1.5), the results show categories of GEN relative clauses are found in the following order: Possession 80 relative clauses, Patient 45, Identity 10, Agent 7, Adverbial 7, Partitive 7 and Reason 5 relative clauses, (Figure 6; for examples of these categories refer to Appendix A).

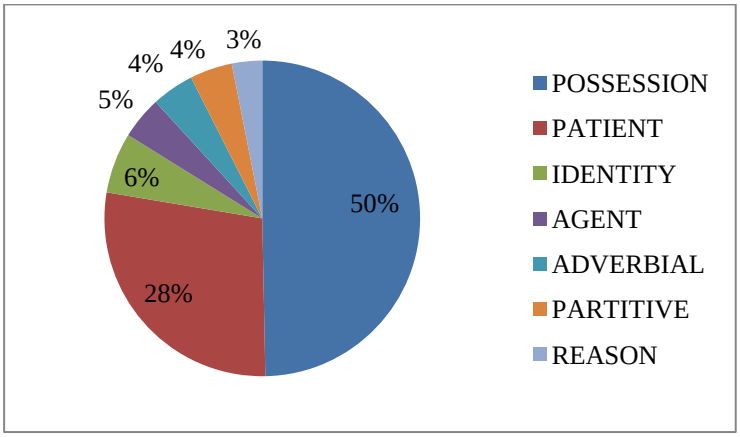


Figure 6: The distribution of GEN relative clauses according to their semantic relationships

5.2.2 The AH

As is shown in the previous sections, TSU+ ISU (1390) are significantly higher than relative clauses in other grammatical positions in the NPAH. Similarly, ISU+DO (1538) are significantly higher than TSU and other relative clauses. At this stage the results of this study conform to both the SPH and the AH. Therefore, to determine which of these hypotheses the results support more, a comparison is made between the SU category, which includes TSU + ISU and the absolutive category, which includes ISU +DO. A model is created in which SU and absolutive are treated as two different

categories, as is shown in Table 10. This model is created to statistically test whether the number of SU relative clauses is significantly higher than the number of absolutive relative clauses and other relative clauses or vice versa.

Table 10: Model 2: Mixed Effect Poisson Regression for the Distribution of Relative Clauses Including Absolutive Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.739	0.138	12.558	<0.001*
ABS	0.122	0.044	2.805	0.005*
DO	-0.613	0.079	-7.738	<0.001*
OBL	-1.128	0.108	-10.414	<0.001*
GEN	-2.155	0.147	-14.636	<0.001*
DEF	1.007	0.034	29.270	<0.001*
INANIM	1.040	0.035	29.978	<0.001*
IA	-1.090	0.211	-5.156	<0.001*
MSA	0.072	0.137	0.527	0.060
NONFIC	0.371	0.137	2.719	0.007*

Model 2 tests five predictors: (1) relative clauses types with the absolutive relative clauses (ABS) as a category, (2) definiteness, (3) animacy, (4) variety and (5) genre; these predictors appear in the first column of the table. The dependent variable is the count of relative clauses of the relevant category. Rows 2-5 of Table 10 show a comparison between SU and other relative clauses. The results in Table 10 suggest that the number of ABS relative clauses is significantly higher than the number of relative clauses in the SU category (*Wald's z*=2.805, *p*=0.005). This result supports the AH, as will be discussed in section 6.3. Results in Table 10 also show that the difference between MSA and CA is not significant, which is not similar to the results shown in Table 9, where the difference between MSA and CA is significant. The difference between the results of these two models might be attributed to the existence of ABS relative clauses in this model, but not in Model 1.

5.2.2.1 Functions of Arabic Relative Clauses

Fox (1987) suggested that the functions that relative clauses have in a text have a great effect on their frequency. Some relative clauses are used to describe/characterise the noun phrase while others are used to relate the head noun phrases to their texts by the use of anchors. The head of the relative clause can be anchored into the text by a noun phrase inside the relative clause which is always known or previously given to the addressee; in other words, noun phrases can only be anchors if they are definite. In TSU relative clauses, DO is the anchor, while in DO relative clauses TSU is the anchor. TSU, which is used inside the DO relative clause, is a better anchor than DO, which is used in the TSU relative clause, because subjects mostly provide given information in contrast with direct objects.

To test Fox's hypothesis, I have examined SU and DO relative clauses that have an anchoring function in Arabic texts. ISU relative clauses are not included in the anchoring test because they are mainly descriptive so they provide characterisation of the head rather than anchoring it to the discourse (Fox, 1987, p. 859). Whether the functions of relative clauses can be used to predict the frequency of relative clauses in the lower positions of the NPAH is a question that needs to be investigated (Song, 2001, p. 241). Therefore, in this study, in addition to TSU and DO relative clauses, I have included the OBL relative clauses. IO relative clauses are not included since there is not enough data for the IO position.

OBL relative clauses have two functions: a characterisation function and an anchoring function, as is shown in examples 83 and 84, respectively. Only 54 OBL relative clauses (11.36%) perform the characterisation function, and 421 OBL relative clauses (88.63%) perform the anchoring function, which will be included in the anchoring texts.

83. wa taḥmilu-hu ila mawāiḍiʿ fī-ha šaǧar
and FEM.carried-him to places in-them trees
muṯmir
fruitful
and (she) carried him to places in which there are fruitful trees.
[CA, 2: 27]

84. al-bait allaḍī huwa sākin fī-hi baitu-nā
the-house REL(3.SG.MAS) he resident in-it house-our
The house in which he lives is our house.
[CA, 1: 311]

In this study, anchors in TSU, DO and OBL relative clauses have been classified, based on the animacy hierarchy, into three categories: first and second person pronoun, third person pronoun, and definite noun. The animacy hierarchy is a good basis since "most discourse is oriented first to the speaker, then to the addressee, and then to the third persons and inanimates along the lines of the hierarchy [the animacy hierarchy] proposed by Dixon (1979: 85)"(Allen & Schroder, 2003, p. 328). The results are shown in Table 11 below:

Table 11: Anchors in Arabic Relative Clauses

Relative clauses	First/Second Pronouns	Third pronouns	Full Noun Phrases
SU	72 (14.09%)	88 (17.22%)	351 (67.51%)
DO	210 (28.64%)	208 (28.37%)	314 (42.97%)
OBL	67 (16.22%)	144 (34.86%)	202 (48.91%)

As can be seen from the data shown in Table 11, full noun phrases functioning as anchors have the highest percentages in the three types of relative clauses. This result differs from Fox's (1987) results as she found that DO relative clauses are anchored to discourse mainly using pronouns. The difference between the results of this study and

Fox's study might be attributed to the fact that this study is based on data from written texts whereas Fox's study is based on data from spoken discourse; this will be discussed further in section 6.3.

A model is created to test the interaction between types of relative clauses and the anchors used, as is shown in Table 12.

Table 12: Model 3: Mixed Effect Poisson Regression for the Interaction between Different Anchors and Types of Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.307	0.240	5.439	<0.001*
3rd pronoun	0.201	0.159	1.262	0.207
Def. NP	1.584	0.130	12.237	<0.001*
DO	1.166	0.167	6.963	<0.001*
OBL	-0.092	0.245	-0.375	0.708
IA	-0.962	0.297	-3.240	0.001*
MSA	0.443	0.210	2.113	0.035 *
NONFIC	0.028	0.210	0.132	0.895
3rd pronoun: DO	-0.210	0.187	-1.126	0.260
Def. NP: DO	-1.185	0.157	-7.538	<0.001*
3rdpronoun:OBL	0.558	0.217	2.565	0.010 *
Def NP:OBL	-0.481	0.192	-2.510	0.012 *

There are four predictors in Model 3: (1) the anchors, which consists of three types (1st and 2nd pronouns, 3rd pronouns and definite noun phrase), (2) relative clauses, which consist of three types (TSU, DO and OBL), (3) variety and (4) genre. Definiteness and animacy are not included as predictors in this model since anchors are all definite, and they are already classified according to the animacy hierarchy.

Rows 3 and 4 of Table 12 confirm the results shown in Table 10, that is, definite noun phrases and third person pronouns are used more than first person pronouns, yet the difference is found significant only between first person pronouns and full noun phrases (*Wald's z*=12.237, *p<0.001*). Rows (6-8) of Table 12 show that there is a significant difference between CA and MSA (*Wald's z*= 2.113, *p*= 0.035) on the one

hand, and CA and IA (*Wald's z*= -3.240, *p*= 0.001) on the other. However, there is no significant difference between the two genres (*Wald's z*= 0.132, *p*= 0.895), which is different from the results in Table 9. The difference between the results of Table 9 and Table 11 might be attributed to the fact that Model 3 in Table 12 includes only relative clauses that have an anchoring function while Model 1 in Table 9 includes all relative clauses in the sample for this study.

The final four rows show that there are some significant interactions between DO and full noun phrase anchors (*Wald's z*= -7.538, *p<0.001*), OBL and third person pronoun anchors (*Wald's z*= -2.565, *p*= 0.010), and between OBL relative clauses and full noun phrase anchors (*Wald's z*= -2.510, *p*= 0.012). These interactions are better shown in Figure 7.

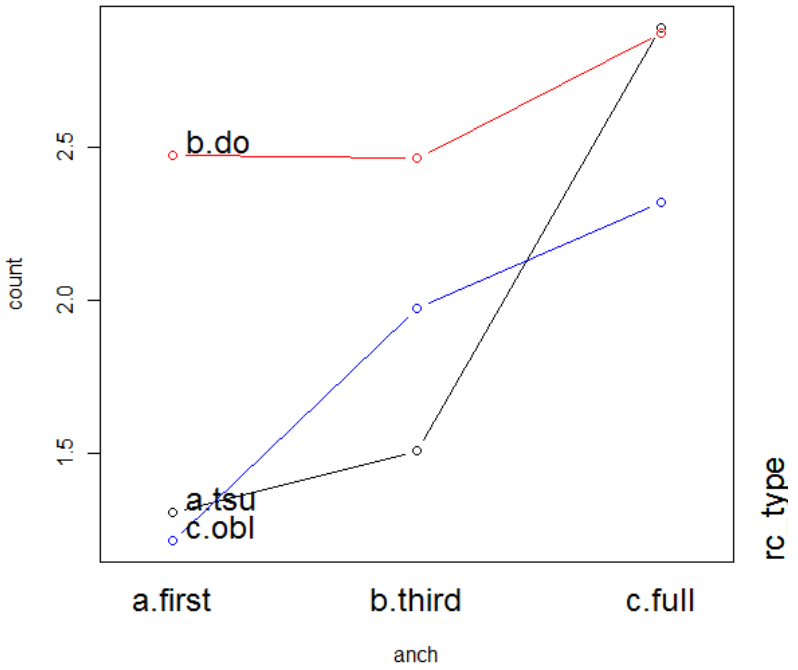


Figure 7: The interaction between anchors and relative clause types

First= first and second personal pronoun, third=third personal pronoun, full=full definite noun phrases

There are nine points in Figure 7.The three on the left show DO, TSU and OBL relative clauses that use first and second person pronouns as anchors. The three points in the middle show DO, TSU and OBL relative clauses that use third person pronouns as anchors. The three points on the right show DO, TSU and OBL relative clauses that use definite noun phrases as anchors. As is shown in Figure 7, the point that represents DO relative clauses using first/second person pronouns anchors is higher than those representing TSU and OBL, which indicates that DO relative clauses have used first and second person pronouns as anchors more than other relative clauses. Figure 7 also shows that relative clauses in the three grammatical positions have a high tendency towards using full noun phrases as anchors.

Separate models for each type of relative clauses are created to test the distribution of anchors in each type of relative clauses (see Tables 4, 5, 6 in Appendix B for the results of TSU, DO and OBL relative clauses, respectively). Results of TSU relative clauses show that there is a greater tendency to use full noun phrases as anchors followed by third person pronouns and then first/second person pronouns; while the difference between first/second person pronouns and full noun phrases is significant (*Wald's z*=12.232, *p<0.001*), the difference between first/second and third person pronouns is not. DO relative clauses also show a great tendency to use noun phrases as anchors with a significant difference between first/second person pronouns and full noun phrases (*Wald's z*=4.471, *p<0.001*). However, unlike TSU relative clauses, their tendency to use first/second person pronouns as anchors is greater than the tendency to use third person pronouns, although the difference between the two is still not significant. This supports the result stated above that DO relative clauses use pronouns as anchors more than TSU relative clauses, and the difference between the two types of relative clauses disappears for full noun phrases. Accordingly, this result supports the

AH which claims that TSUs in DO relative clauses provide better anchors than DOs in TSU relative clauses do.

OBL relative clauses also show a tendency to use full noun phrases as anchors and the difference between full noun phrases and first/second person pronoun anchors is significant (*Wald's z*=7.819, *p<0.001*). Third person pronouns are used as anchors in OBL relative clauses significantly more than first/second person pronouns (*Wald's z*=5.115, *p<0.001*); OBL relative clauses are the only relative clauses which show a significant difference between these two types of anchor (i.e. first/second person pronouns and third person pronouns). A comparison between TSU relative clauses and OBL relative clauses shows that TSU relative clauses use third and first/second person pronouns as anchors less than OBL relative clauses, and they use full noun phrases more than OBL relative clauses. This result indicates that OBL relative clauses fit better than TSU relative clauses to the anchoring function. The possible motivation for these results will be discussed in 6.3.

5.3 Definiteness

There are a number of factors that are considered in this study to determine their effects on the distribution of relative clauses in Arabic texts. Definiteness is one of these factors; it is worth studying due to the fact that definite relative clauses are different from indefinite relative clauses not only in the definite status of the head, but also due to the fact that relative markers are used in definite relative clauses only (refer to section 3.3.1.1). This section presents the results of classifying Arabic relative clauses according to definiteness.

Of all Arabic relative clauses included in the study, 2022 relative clauses are definite, and 763 are indefinite. Definite relative clauses which make up to 72% outrank indefinite relative clauses. The difference between definite and indefinite relative clauses is highly significant (*Wald's z*=23.105, *p<0.001*), as is shown in Table 9.

In regards to the order of frequency of definite and indefinite relative clauses, the order of definite relative clauses shows that DO and ISU relative clauses are significantly more frequent than TSU relative clauses, and there is no significant difference between TSU and OBL. On the other hand, the order of indefinite relative clauses shows that there is no significant difference among ISU, DO and TSU, which are more frequent than OBL relative clauses (see Tables 7-10 in Appendix B), as is shown in Figure 8.

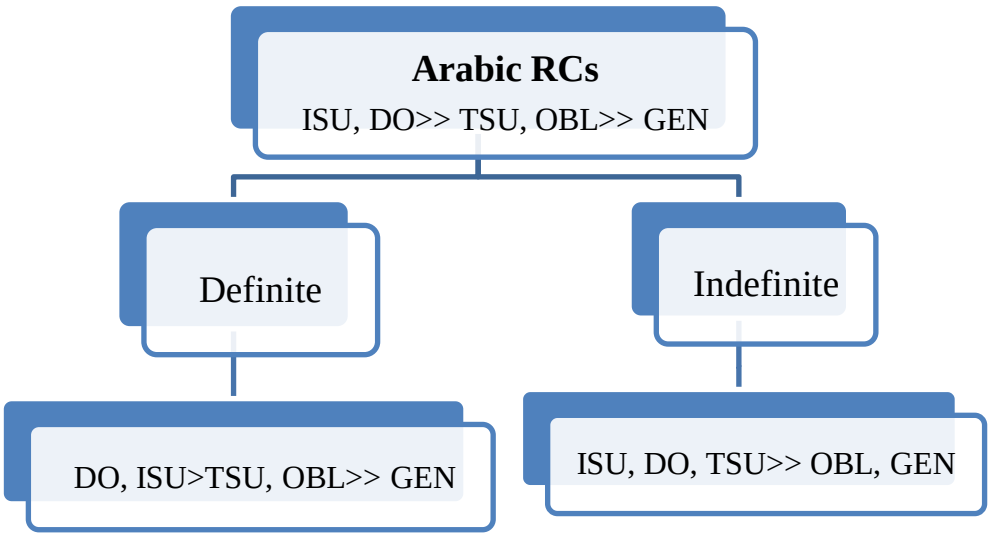


Figure 8: The distribution of definite and indefinite relative clauses in Arabic texts, where (,)[24] indicates that the difference between the two positions is not significant; (>)indicates approaching significance; (>>) indicates significant

[24] These symbols are used for the purpose of illustrating the differences as far as statistical significance in the results of this study only, and not to be confused with (>) used in the original NPAH.

The AH is manifested in both distributions although it is manifested more in definite relative clauses. The reason behind that is that, in indefinite relative clauses, the DO and TSU relative clauses are found to have the same percentage and SU relative clause (TSU+ISU) and ABS relative clauses (ISU+DO) have the same percentage (52.42%), whereas, in definite relative clauses, the percentage of ABS relative clauses (56.28%) is higher than the percentage of SU relative clauses (48.96%). Therefore, excluding indefinite relative clauses from the sample of this study would not make the distribution of relative clauses in Arabic texts conform to the order of accessibility suggested by the NPAH.

As is shown in Figure 8, tendencies to have definite or indefinite relative clauses are different among different grammatical positions. A model is created to test possible significant interactions between definiteness and relative clauses in different positions, as is shown in Table 13.

Table 13: Model 4: Mixed Effect Poisson Regression for the Interaction between Definiteness and Distribution of Relative Clauses in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.341	0.141	9.518	<0.001*
TSU	-0.402	0.204	-1.971	0.049 *
DO	-0.297	0.157	-1.890	0.059 .
OBL	-0.772	0.162	-4.776	<0.001*
GEN	-1.123	0.194	-5.800	<0.001*
DEF	0.882	0.078	11.264	< 0.001 *
INANIM	1.059	0.043	24.379	< 0.001 *
IA	-1.273	0.223	-5.698	<0.001*
MSA	-0.279	0.127	-2.205	0.027 *
NONFIC	0.542	0.127	4.281	<0.001*
TSU:DEF	0.057	0.120	0.472	0.637
DO:DEF	0.351	0.117	2.994	0.003 *
OBL:DEF	0.306	0.134	2.286	0.022 *
GEN:DEF	-0.718	0.178	-4.038	<0.001*

Model 4 has five predictors: (1) relative clauses in (ISU, TSU, DO, OBL, GEN) grammatical positions, (2) definiteness, (3) animacy, (4) genre and (5) variety. The comparison between ISU relative clauses and other relative clauses, which is displayed in rows 2-5, shows that the differences between ISU and other relative clauses are significant except in DO, which is approaching significance, which is slightly different from the results in Table 9, where the difference between DO and ISU is not significant. This slight change takes place due to the interaction between definiteness and relative clause type, which is not counted in Model 1 in Table 9.

The final five rows in Table 12 show the interaction between relative clauses in the tested grammatical positions and definiteness. The results show there is a significant interaction between relative clause type and definiteness, which can be interpreted in Figure 9.

Figure 9 shows five types of relative clauses (ISU, TSU, DO, OBL, GEN). Points on the red graph show the frequencies of definite relative clauses while the points on the black graph show the frequencies of indefinite relative clauses; relative clauses in different grammatical positions are compared to relative clauses in the ISU position. The count axis shows how frequently relative clauses are used in these grammatical positions. The count for definite clauses is higher in all positions except for GEN. These visual differences are confirmed in pairwise comparisons (see Tables 11-15 in Appendix B for TSU, ISU, DO, OBL and GEN, respectively).

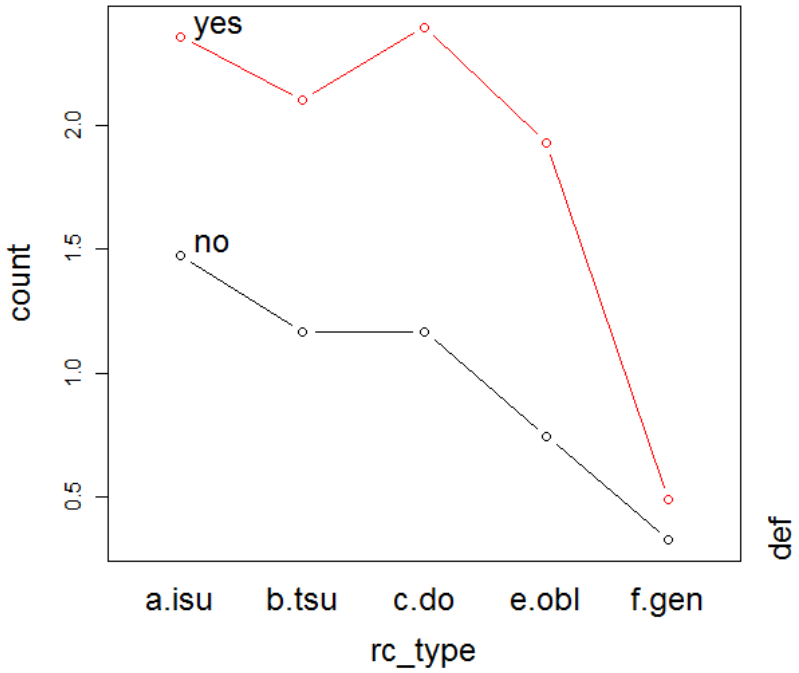


Figure 9: The interaction between the grammatical position of the relativised noun phrase and definiteness, where 'yes' represents definite and 'no' represents indefinite

Results show that the relativised noun phrases in TSU, ISU, DO and OBL positions show a positive tendency to be definite and the difference between definite and indefinite relative clauses is significant in these positions: TSU (*Wald's z*=10.336, *p<0.001*), ISU (*Wald's z*=11.263, *p<0.001*) DO (*Wald's z*=14.096, *p<0.001*) and OBL (*Wald's z*=10.912, *p<0.001*). On the other hand, there is no significant difference between definite relative clauses and indefinite relative clauses in the GEN position (*Wald's z*=1.024, *p*=0.3058). Therefore, all relative clauses have significant tendencies towards identifying definite heads except GEN relative clauses. These results show an agreement with Gordon and Hendrick's (2005) results, where the authors found that there is a tendency of ISU and DO relative clauses to have definite heads. On the other hand, other discourse studies (e.g. Fox, 1987; Fox & Thompson, 1990; Hogbin & Song,

2007) have suggested that ISU relative clauses tend to have indefinite heads as these relative clauses have the characterisation function. The difference between the results of the present study and results of previous studies will be discussed further in section 6.6.1.

Definiteness of the head of a relative clause can be predicted based on the head's position in the main clause. This has been manifested in previous studies such as Fox and Thompson (1990) as the existential heads tend to be indefinite, which increased the percentage of indefinite relative clauses in the ISU position in that study. In the following section, we will present the results of the grammatical position of the head in the main clause in relation to the distribution of relative clauses in discourse on the one hand and definiteness of relative clauses on the other.

5.4 The Position of the Head in the Main Clause

Relative clauses in this study are classified, according to the position of their heads in the main clause, into SU, SU(TOPIC[25]), DO, Subject complement (SUCOM), OBL, GEN and OTHERS. The category of 'OTHERS' includes existential (EXIS), adverbs (ADV), cognate accusative[26] (CACC), objects of comparison (OCOMP), indirect objects (IO), and object complements (OCOM), as is shown in Figure 1, where head NPs have a high tendency towards occupying the OBL position with a percentage of 31%.

[25] Subjects are classified into SU and SU(TOPIC); the latter includes subjects which precede their verbs, thus are in the topic position. This classification arises from the different structures that can be found in Arabic. Arabic sentences can be found as VS(O) or SV(O) in addition to the verbless sentences where the subjects are almost always in the topic position.

[26] "The cognate accusative emphasizes or intensifies a statement by using a verbal noun derived from the main verb or predicate" (Ryding, 2005, p. 83)

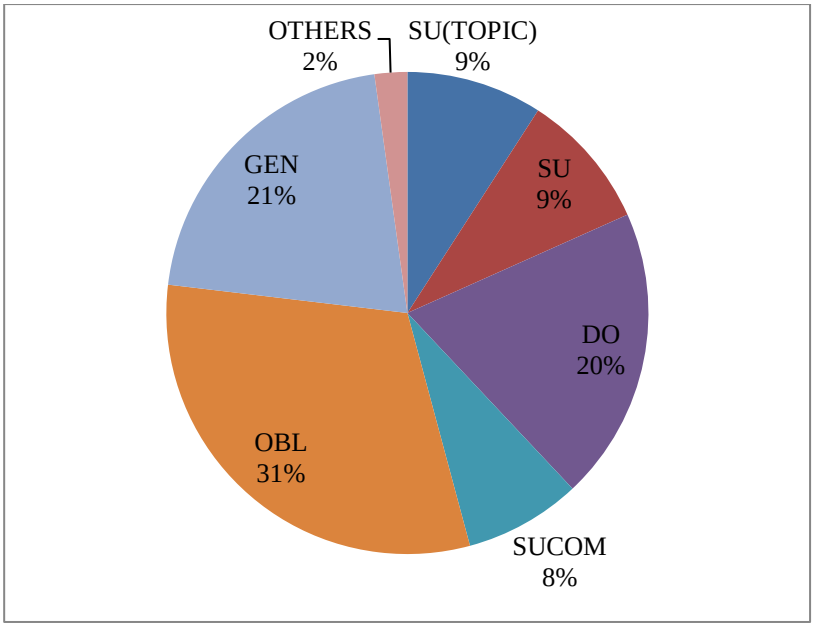


Figure 10: The distribution of relative clauses according to the grammatical position of the head.

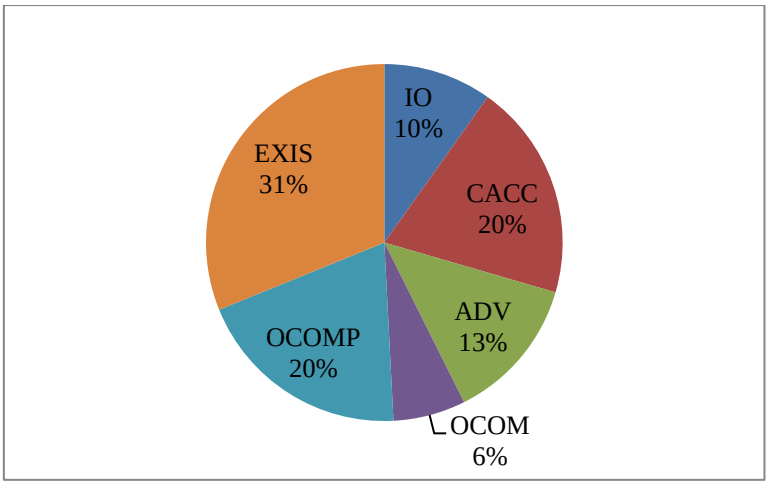


Figure 11: The classification of the category of 'OTHERS' according to their percentages within this category

Tendency of head noun phrases to occupy OBL is found to be statistically significant in ISU, DO and OBL relative clauses. In TSU relative clauses the difference between OBL and other grammatical positions of the head is significant except in the

DO position. The results show that the difference between OBL heads and DO heads in TSU relative clauses is approaching significance (*Wald's z*= -1.898 , *p*=0.058). On the other hand, the results for GEN relative clauses show that the differences between OBL heads and OTHERS (*Wald's z*= -3.983, *p<0.001*), SUCOM (*Wald's z*= -3.320, *p*=0.001), and SU(TOPIC) (*Wald's z*= -4.090, *p<0.001*) are the only significant differences found, although the OBL heads are found more frequently than other heads (see Tables 16-20 in Appendix B).

As is explained in section 4.4.2, different groups or categories of relative clauses can be detected based on the head's position in the main clause and the relative clause. For example, if the head occupies a subject position in the main clause and a DO position in the relative clause, then this combination will be written as SU-DO. Accordingly, relative clauses in this study are found to have different combinations. For a comprehensive list of these combinations and examples, refer to Appendix C. The numbers of relative clauses in these combinations are shown in Table 14.

As is shown in Table 14, the number of relative clauses in the category SU-ISU is 100, and the number of SU-TSU relative clauses is 53, and so on in other categories. To find the effect of the grammatical position on the distribution of relative clauses in discourse, it is important to find the tendency for the head in each type of relative clause to occupy a certain grammatical position. The results shown in Table 13 emphasize the strong tendency in all kinds of relative clauses to have OBL heads (IO and OCOMP relative clauses are removed from the table because of their small numbers).

Table 14: The Distribution of Heads NPs among Grammatical Positions in the Main and Relative Clauses

	Positions of head NPs in the Relative Clauses				
Positions of head NPs in the main clause	ISU(%)	TSU(%)	DO(%)	OBL(%)	GEN(%)
SU	100(12.67)	53(8.82)	47(6.30)	26(5.51)	26(16.15)
SU(TOPIC)	67(8.49)	66(10.98)	78(10.46)	34(7.20)	7(4.35)
DO	165(20.91)	118(19.80)	158(21.18)	72(15.47)	31(19.26)
OBL	236(29.91)	176(29.29)	207(27.75)	207(43.86)	39(24.22)
GEN	164(20.79)	122(20.30)	173(23.19)	86(18.22)	37(22.98)
SUCOM	43(5.45)	48(7.99)	76(10.19)	37(7.84)	13(8.08)
OTHERS	14(1.77)	17(2.83)	7(0.94)	9(1.91)	8(4.97)

To determine the relationship between grammatical positions of the heads of relative clauses and the distribution of these relative clauses, it is also important to determine whether head NPs have different tendencies to occupy certain grammatical positions rather than others in relative clauses. A number of models are created to test the distributions of relative clauses whose heads occupy OBL, GEN, DO, SU(TOPIC), SU, SUCOM and OTHERS in the main clause (see Tables 21-27 in Appendix B). The distributions are shown in Table15 where relative clauses appear in a descending order.

Table 15: NPs in Different Grammatical Positions and Relative Clauses Identifying Them

Heads	Order of relative clauses
OBL	ISU, DO, OBL, TSU, GEN
GEN	DO, ISU, TSU, OBL, GEN
DO	ISU, DO, TSU, OBL, GEN
SU	ISU, TSU, DO, OBL, GEN
SU(TOPIC)	DO, ISU, TSU, OBL, GEN
SUCOM	DO, TSU, ISU, OBL, GEN
OTHERS	TSU, OBL, ISU, DO, GEN

As is shown in Table 15, relative clauses which identify NPs occupying different grammatical positions in the main clauses have different orders, which emphasises further the importance of the grammatical position of the head in the main clause.

As discussed in 3.3.1.2.2, Sheldon (1974) suggests that relative clauses are more accessible to comprehension and acquisition if the grammatical position of the head in the main clause parallels with its position in the relative clause, which he calls the 'Parallel Hypothesis'. As is shown in Table 14, the Parallel Hypothesis is not confirmed by the results of this study.

Cases where the grammatical positions in the relative clauses are parallel with the grammatical positions in the main clauses do not show high frequencies except in OBL relative clauses. OBL-OBL relative clauses (207) are more frequent than other OBL relative clauses such as SU-OBL, DO-OBL etc. Nonetheless, the most frequent combination found in this study is OBL-ISU where the grammatical position of the head in the main clause does not match its grammatical position in the relative clause. A discussion of the parallel hypothesis in relation to the data of this study is presented in 6.6.2.

Definiteness of the noun phrase is associated with its grammatical position in the sentence. Noun phrases tend to be in the definite case when they occupy certain grammatical positions, such as SU, rather than the others (Givón, 1979; Lyons, 1999). A model is created to test the interaction between the grammatical positions of the heads of relative clauses and the definiteness of relative clauses, as is shown in Table 16.

Table 16: Model 5: The Interaction Between the Grammatical Position of the Head in the Main Clause and its Definiteness

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.432	0.141	10.150	< 0.001 *
TSU	-0.411	0.196	-2.101	0.036*
DO	-0.080	0.148	-0.543	0.587
OBL	-0.560	0.135	-4.156	< 0.001 *
GEN	-1.537	0.160	-9.618	< 0.001 *
GEN HEADS[27]	-0.810	0.120	-6.769	< 0.001 *
OBL HEADS	-0.243	0.100	-2.428	0.015 *
OTHERS HEADS	-1.575	0.160	-9.824	< 0.001 *
SUCOM HEADS	-1.298	0.143	-9.054	< 0.001 *
SU HEADS	-0.542	0.110	-4.951	< 0.001 *
TOPIC HEADS	-3.346	0.360	-9.297	< 0.001 *
DEF	0.339	0.089	3.830	0.000*
NONFIC	0.515	0.132	3.916	< 0.001 *
IA	-1.250	0.223	-5.611	< 0.001 *
MSA	-0.229	0.132	-1.740	0.082.
GEN HEADS:DEF	1.130	0.142	7.975	< 0.001 *
OBL HEADS:DEF	0.926	0.123	7.546	< 0.001 *
OTHERS HEADS:DEF	-1.641	0.335	-4.893	< 0.001 *
SUCOM HEADS:DEF	0.641	0.174	3.678	0.000*
SU HEADS:DEF	-0.476	0.157	-3.036	0.002*
TOPIC HEADS:DEF	3.048	0.371	8.224	< 0.001 *

Model 5 tests five predictors: (1) relative clause types, (2) grammatical positions of the heads of relative clauses, (3) definiteness, (4) variety and (5) genre. Animacy is not included in this model because the main intention was to test the interaction between definiteness and grammatical positions of the heads in the main clauses. Therefore, definiteness and the grammatical position of the head are added to types of relative clauses, variety and genre, which are the three factors that this study originally set out to test.

[27] A GEN HEAD= the head of the relative clause occupies the GEN position in the main clause, so it is different from a GEN relative clause. This is applied to other types of heads such as OBL HEADS etc...

As is shown in Table 16, the results in Model 5 are similar to the results obtained in Model 1 except that the difference between CA and MSA is not significant, but approaching significance in Model 5. This change takes place because of the interaction between the grammatical positions and definiteness.

Regarding the grammatical position of the head, Model 4 suggests that relative clauses whose DO heads are used significantly more than other relative clauses. These results differ from the results suggested by the percentages of relative clauses (see Figure 10), which indicates that relative clauses with OBL heads are the most frequent relative clauses. This difference is also due to the interaction, which is tested in Model 4; this is confirmed by another model that tests similar predictors but without considering the interaction (see Table 28 in Appendix B), which shows that OBL heads are found significantly more than DO heads (*Wald's z*= 6.735, *p<0.001*).

The last six rows of Table 16 show the interaction between the grammatical positions of the heads of relative clauses and definiteness. The results indicate that the interactions between definiteness and all the grammatical positions tested are significant. These interactions are better shown in Figure 12.

There are seven points in Figure 12, which represent the grammatical positions of the heads in the main clause (TOPIC, SU, DO, OBL, GEN, SUCOM, OTHERS). Heads of relative clauses have the tendency to be definite in all grammatical positions except in two positions, the "OTHERS" and the "SU" positions. The OBL position shows the highest tendency to be definite followed by GEN, DO, TOPIC and SUCOM, respectively. Thus, the position of the head in the main clause has an effect on definiteness, and this effect is statistically significant.

The question of how the interaction between the grammatical position of the head and its definiteness has contributed to the frequent use of definite relative clauses in Arabic texts will be discussed in chapter 6.6.1.

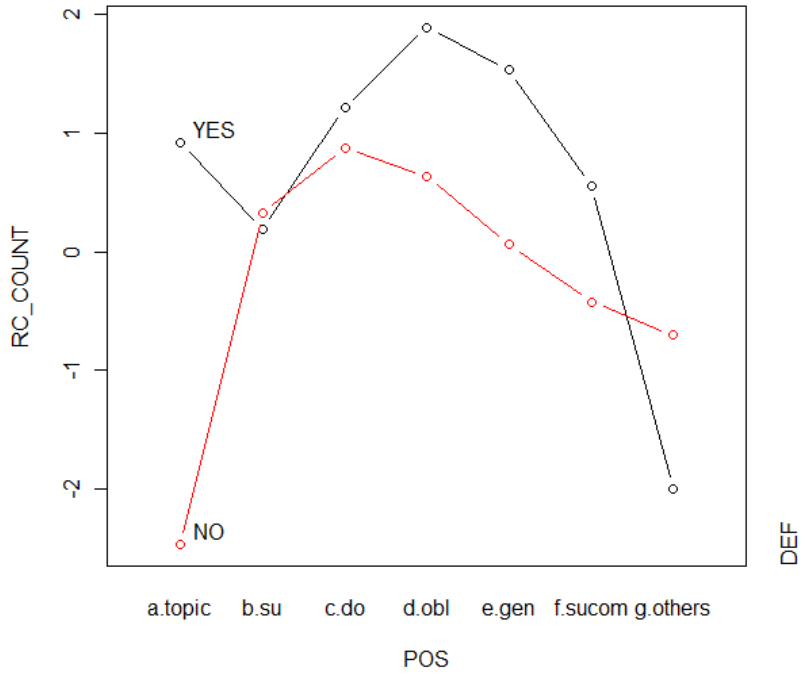


Figure 12: The interaction between the grammatical position of the head in the main clause and its definiteness, where 'YES' represents definite and 'NO' represents indefinite

5.5 Animacy

Animacy is believed to have a role in accessibility to relativisation (Hawkins, 2007; Jeon & Kim, 2007), and it is also considered to be one of the factors that affects the distribution of relative clauses in discourse (Fox & Thompson, 1990). Therefore, this section presents the results of animacy in this study and its relation to definiteness and the distribution of relative clauses in discourse.

Relative clauses whose heads refer to inanimate entities form the highest percentage (74%) of all relative clauses. The second category consists of those relative clauses whose heads refer to humans (23%). Finally, come relative clauses referring to nonhuman animates (3%).

Due to the small percentage of relative clauses that refer to nonhuman animates, they are not included as a separate category in the statistical models of this study. Instead, they are added to relative clauses referring to human referents and together they form the 'animate' category. As is shown in Model 1, the difference between relative clauses that refer to animate referents are significantly less than those refer to inanimate referents (*Wald's z*= 24.379, $p< 0.001$).

The distribution of relative clauses referring to animate referents is different from the distribution of relative clauses referring to inanimate referents since and both distributions are different from the overall distribution of relative clauses in this study, as is shown in Figure 13. With relative clauses that refer to animate, the number of ISU and TSU (thus SU) relative clauses is found significantly higher than the number of DO and other types of relative clauses. On the other hand, relative clauses that refer to inanimate referents, TSU relative clauses are used significantly less frequently than DO, ISU and OBL relative clauses (refer to Tables 29-33 in Appendix B).

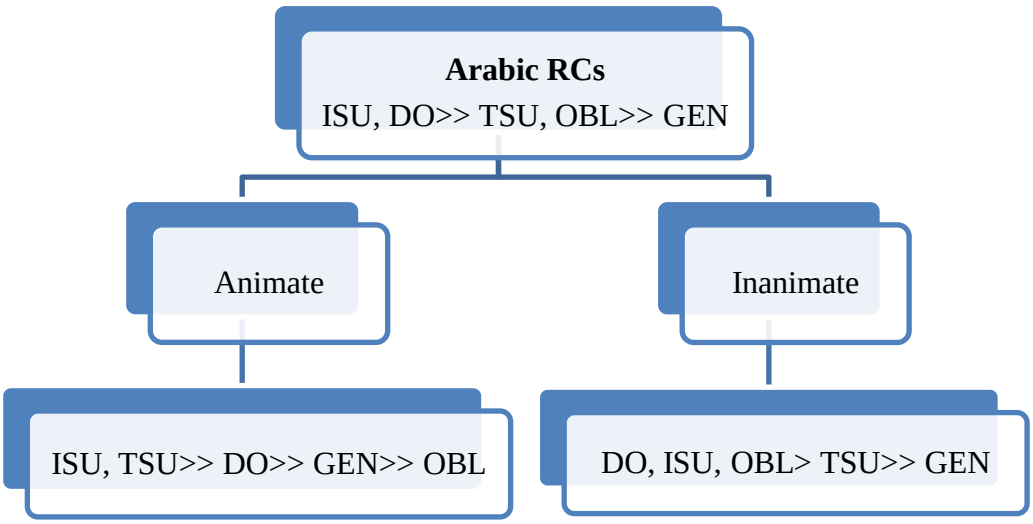


Figure 13: The distribution of relative clauses whose heads refer to animate and inanimate referents

The distribution of relative clauses shown in Figure 13 suggests that there could be a tendency for some relative clauses in certain grammatical positions in the NPAH to identify heads referring to animate referents more than others do. Model 6 tests the possible interaction between animacy and relative clauses in different positions in the NPAH.

Table 17: Model 6: The Interaction between Animacy and Relative Clauses in the NPAH

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.679	0.137	12.274	< 0.001*
TSU	-0.437	0.198	-2.203	0.028 *
DO	-1.048	0.165	-6.365	< 0.001*
OBL	-2.399	0.225	-10.671	< 0.001*
GEN	-1.691	0.202	-8.356	< 0.001*
INANIM	0.451	0.073	6.173	< 0.001*
DEF	0.987	0.043	23.105	< 0.001*
NONFIC	0.542	0.121	4.281	< 0.001*
IA	-1.273	0.223	-5.698	< 0.001*
MSA	-0.279	0.127	-2.205	0.0274 *
TSU:INANIM	-0.173	0.336	-0.515	0.606
DO:INANIM	1.352	0.128	10.582	< 0.001*
OBL:INANIM	2.282	0.205	11.126	< 0.001*
GEN:INANIM	0.186	0.183	1.018	0.309

There are five predictors in Model 6: (1) relative clauses in five positions on the NPAH (ISU, TSU, DO, OBL, GEN), (2) animacy (animate and inanimate), (3) definiteness, (4) variety and (5) genre, which are the same predictors included in Model 1. Results in Model 6 are different from results of Model 1 in that the difference between DO relative clauses and ISU relative clauses in Model 6 is significant (*Wald's* z=-6.365, p=0.001), which is not the case in Model 1. The difference between TSU relative clauses and ISU relative clauses is approaching significance in Model 1, and significant in Model 6 (*Wald's* z=-2.203, p=0.028). These differences between the two models might be attributed to the effect of interaction between relative clauses and animacy, which is tested in Model 6.

As is shown in the last four rows of Table 16, there are significant interactions between heads refer to inanimate referents and some positions such as DO (*Wald's* z=10.582 , $p<0.001$) and OBL (*Wald's* z= 11.126, $p<0.001$). These interactions are better shown in Figure 14.

Each graph in Figure 14 represents a category of animacy: relative clauses modifying inanimate referents are represented by the red line, and relative clauses modifying animate referents are represented by the black line. In each of these graphs, there are five points which represent the relative clauses, ISU, TSU, DO, OBL and GEN. As is shown in Figure 14, while DO and OBL relative clauses have the highest tendency to modify heads with inanimate referents, ISU relative clauses have the highest tendency to modify animate referents. Similarly, GEN relative clauses have a higher tendency to occur with heads refer to animate referents than OBL. On the other hand, the tendency of GEN relative clauses to occur with inanimate referents is low compared to OBL. This might explain the difference in the order of frequency of

relative clauses referring to animate referents and the order of frequency of relative clauses referring to inanimate referents, which will be discussed in 6.6.3.

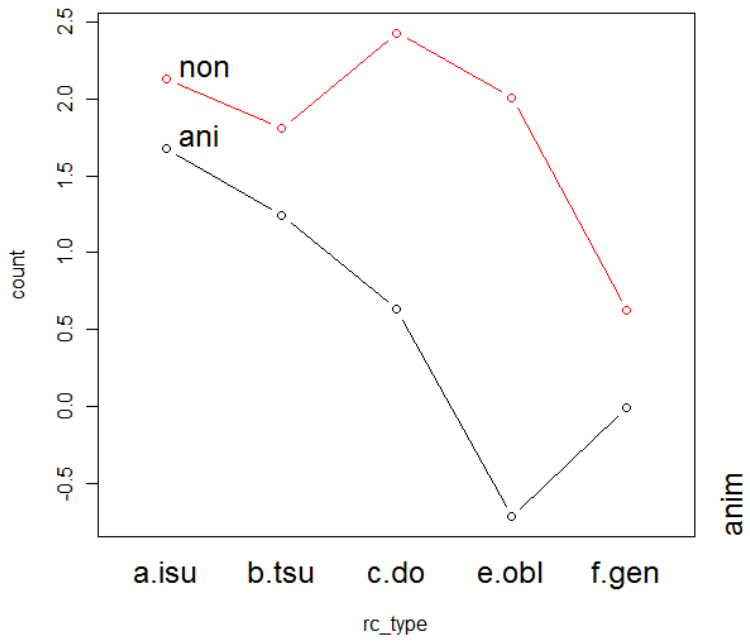


Figure 14: The interaction between relative clauses in different grammatical positions on the NPAH and animacy, where 'ani' represents animate and 'non' represents inanimate

Animacy has also been found to affect the definiteness of noun phrases since noun phrases that refer to human referents tend to be definite (Givón, 1979; Lyons, 1999). However, the results of the model, which was created to test the possible interaction between animacy and definiteness, show that there is a significant positive interaction between animacy and definiteness (*Wald's z*=5.064, *p<0.00;* see Table 34 in Appendix B). This suggests that the heads refer to inanimate referents are more likely to be definite than heads refer to animate referents.

The following sections will examine the effect of construction-external variables, genre and variety, on the distribution of relative clauses in discourse. I will also

examine whether there is any relation between these two factors, variety and genre, and the construction-internal variables that have been presented in the previous sections, which are definiteness, the position of the head in the main clause and animacy.

5.6 Variety

Three varieties of Arabic are included in this study, CA, MSA, IA. In this section, I will analyse the impact of variety across all data. Then, I will examine each variety separately and study the different predictors within that variety; CA will be presented in 5.6.1, MSA will be presented in 5.6.2, and IA will be presented in 5.6.3.

As is shown in Table 1, the differences are significant between CA and the other two varieties, MSA (*Wald's z*=-2.206, *p*=0.027) and IA (*Wald's z*=-5.698, *p<0.001*). This result remains consistent in the other models except in Model 2 in Table 10, where the difference is not significant (*Wald's z*=0.527, *p*=0.060), and in Model 5 in Table 16, where the difference between MSA and CA is approaching significance instead of significant (*Wald's z*=-1.740, *p*=0.082) (see Tables 10, 13, 16 and 17).

To study the influence of variety on the distribution of relative clauses in the text, a model is created to test the interaction between variety and the distribution of relative clauses, as is shown in Table 18.

Table 18: Model 7: Mixed Effect Poisson Regression for the Interaction between Relative Clauses and Varieties

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.197	0.141	8.501	<0.001 *
TSU	-0.959	0.188	-5.103	<0.001 *
DO	-0.395	0.159	-2.489	0.013 *
OBL	-0.628	0.192	-3.263	0.001*
GEN	-1.678	0.244	-6.885	<0.001 *
IA	-1.364	0.258	-5.279	<0.001 *
MSA	-0.085	0.165	-0.512	0.609
NONFIC	0.535	0.126	4.265	<0.001 *
DEF	0.987	0.043	23.104	<0.001 *
INANIM	1.059	0.043	24.378	<0.001 *
TSU:IA	0.894	0.377	2.369	0.018 *
DO:IA	0.784	0.328	2.387	0.017 *
OBL:IA	0.077	0.411	0.187	0.852
GEN:IA	0.559	0.502	1.114	0.265
TSU:MSA	1.080	0.258	4.183	<0.001*
DO:MSA	0.574	0.222	2.588	0.010 *
OBL:MSA	0.148	0.272	0.545	0.586
GEN:MSA	0.148	0.345	0.141	0.888

Model 7 tests the same set of predictors tested in Model 1, which are (1) relative clauses in five grammatical positions, ISU, TSU, DO, OBL, GEN, (2) variety, (3) genre (4) definiteness, and (5) animacy, and it also tests the interaction between types of relative clauses and variety, which is not tested in Model 1. The order of frequency which appears in Table 18 is similar to that in Model 1 except that the difference between the ISU and DO relative clauses is significant (*Wald's z*=-2.489, *p*=0.01282) in Table 18. In relation to varieties, there is a significant difference between IA and CA (*Wald's z*=-5.279, <0.001). On the other hand the difference between MSA and CA does not appear to be significant, which is not the same result obtained in Model 1, where the difference between MSA and CA is significant. The difference between these two results might be attributed to the effect of the interaction between variety and the distribution of relative clauses.

The final eight rows show the results of the interaction between variety and relative clauses. There are significant interactions between TSU and IA (*Wald's* z=2.369, p=0.018), and TSU and MSA (*Wald's* z=4.183, <0.001). There are also significant interactions between DO and IA (*Wald's* z=2.387, p=0.017), DO and MSA (*Wald's* z=2.588, p=0.010). On the other hand, there are no significant interactions between varieties (IA and MSA) and relative clause (see Table 35 Appendix B). The results suggest that the variety in which relative clauses are written plays an important role in deciding the order of the frequency of relative clauses in the upper grammatical positions (ISU, TSU, DO). The interactions are better shown in Figure 15.

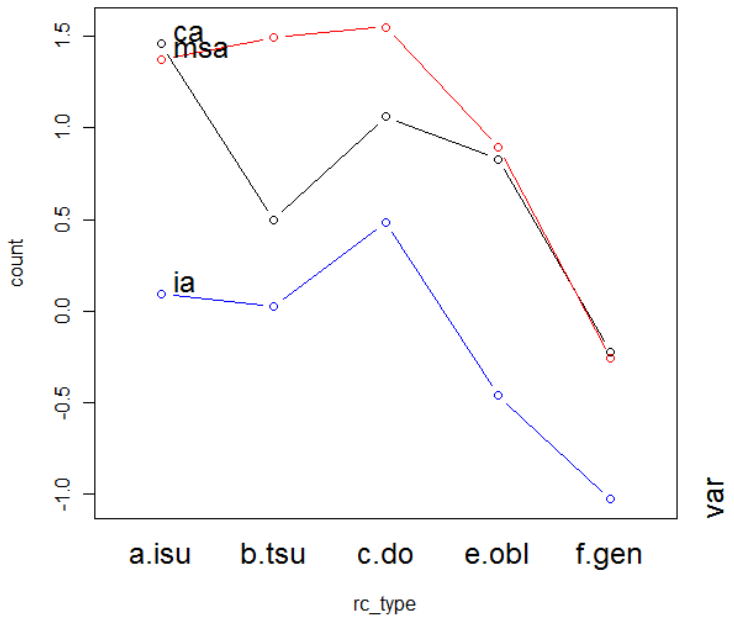


Figure 15: The interaction between relative clauses and varieties

There are three graphs in Figure 15. Each line represents a variety; for each line, there are five points which represent the five grammatical positions tested in this model (ISU, TSU, DO, OBL, GEN). The count axis shows the frequency of relative clauses in

each variety. IA (blue) relative clauses are less than the CA (black) and MSA (red). CA and MSA are very close in the ISU, OBL and GEN positions and only differ significantly in the TSU and DO positions. Similarly, IA differs significantly from CA in the three upper positions, ISU, TSU and DO, but there is a slight difference between IA and MSA in the TSU position which is not significant as is confirmed by the results in Table 35. The differences among varieties will be discussed in relation to the accessibility and the diglossic nature of the Arabic language[28] in section 6.4.1.

In the next sections, I will present the results of each variety in relation to the distribution of relative clauses, definiteness, the position of the head in the main clause and animacy.

5.6.1 CA

The CA texts reveal 1166 relative clauses, which makes up to 41.86% of all relative clauses in the Arabic corpus. As is shown in Table 36 (in Appendix B), the differences between ISU relative clauses and relative clauses in other grammatical positions are significant in CA texts. As is shown in Figure 16, the CA line (in black) has the highest point in the ISU position followed by DO, OBL, where there is a slight non-significant difference between the two positions (see Table 37 in Appendix B), and then comes the TSU position, which is followed by GEN.

[28] The intention of introducing diglossia in section 3.5 is to explain the situation of Arabic to the reader and to focus on the importance of studying the NPAH across varieties of Arabic. Data collection was not set for studying diglossia, yet the results could have implications for that. Therefore, diglossia is discussed in relation to the use of relative clauses in Arabic texts and the NPAH. This limited discussion does not match the importance of diglossia in the literature, yet it shows that this study can have a sociolinguistic dimension that could be followed and enhanced in future studies. It is also important to assert that there is no intention to address code switching and bilingualism theories, although they are related to diglossia, because that is beyond the focus of this study.

To test the difference between SU relative clauses (TSU+ISU) and ABS relative clauses (DO+ISU), a model was created in which TSU and ISU relative clauses were put under SU, and ABS was included as a type of relative clause (see Table 38 in Appendix B). The results show that the ABS is significantly more frequent than the SU (*Wald's z*= 3.180, *p*= 0.002) in CA texts.

Definite relative clauses are more frequent than indefinite relative clauses in CA texts and the difference between the two kinds is highly significant (*Wald's z*=6.000, *p<0.001*), as is shown in Table 36 in Appendix B. This result is in agreement with the result in Table 1, which indicates that definite relative clauses occur significantly more frequently than indefinite relative clauses in the overall data as well as in CA. The order of frequency of relative clauses with definite heads is different from the order of frequency of relative clauses which identify indefinite heads (see Tables 39-44 in Appendix B). The overall distribution of CA relative clauses is found to be different from the distribution of definite relative clauses and that of indefinite relative clauses. Definite relative clauses show a significant difference between the GEN position and other positions. On the other hand, there are not any significant differences among the grammatical positions in the distribution of indefinite relative clauses in CA.

Regarding animacy, most CA relative clauses (68%) modify heads referring to inanimate referents followed by those that describe human referents (28%) and those that describe nonhuman animate referents (4%) come last. Relative clauses that modify inanimate referents are significantly higher than relative clauses referring to animate (*Wald's z*=12.310, *p*< 0.001; see Table 38 in Appendix B). The distribution of relative clauses whose heads refer to animate shows that the frequencies of ISU and TSU relative clauses are significantly higher than other relative clauses. On the other hand,

TSU relative clauses comes fourth, in the distribution of relative clauses whose heads refer to inanimate, after DO, ISU and OBL(see Tables 45-49 in Appendix B).

Nonfiction relative clauses make up (52%) of CA data. The difference between the frequency of fiction and nonfiction relative clauses in CA texts is also significant (*Wald's* z=4.768, p<0.001; see Table 38 in Appendix B). The overall order of CA relative clauses is reflected neither in CA fiction nor in CA nonfiction (see Tables 51-54 in Appendix B). OBL with DO relative clauses come in the second place after ISU in the overall distribution of CA, but they, OBL relative clauses, are used significantly less than DO, ISU and TSU in fiction texts; and used significantly less than ISU, DO but higher than TSU in nonfiction texts, as shown in Figure 16, below.

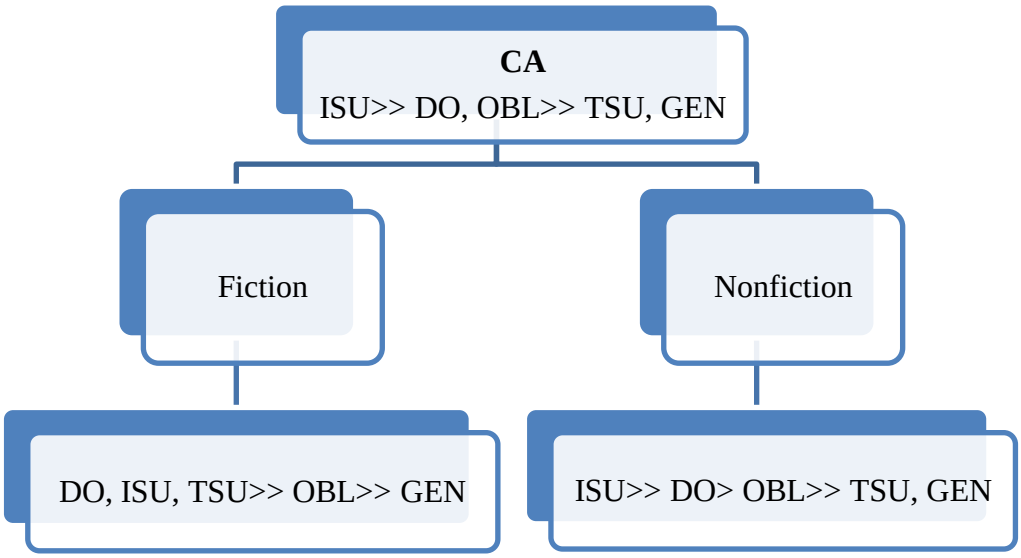


Figure 16: The distribution of CA relative clauses in fiction and nonfiction texts

Most heads of CA relative clauses occupy the OBL position in the main clauses with a percentage of 31% followed by DO, SU and GEN, SUCOM, SU(TOPIC),

OTHERS (which includes: CACC, OCOMP, ADV, IO, OCOM), respectively, as is shown in Figure 17.

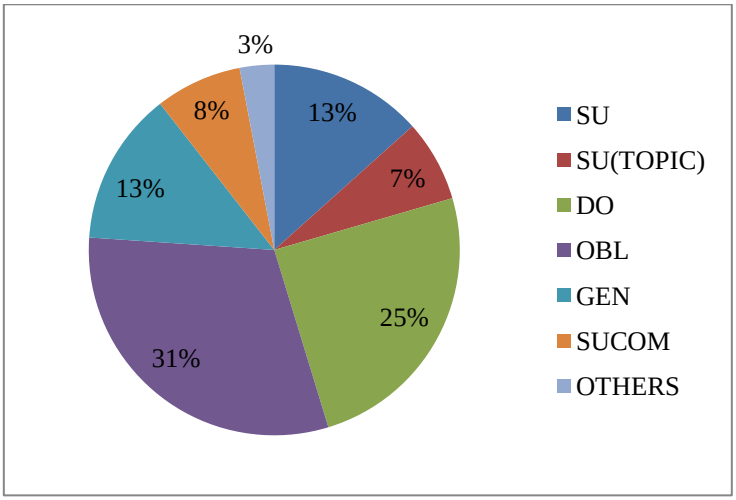


Figure 17: The classification of CA relative clauses according to their head's position in the main clause

The OBL position is significantly higher than other grammatical positions, as is shown in Table 55 in Appendix B. Most grammatical positions, particularly those with the highest percentages, have a preference to be identified by either ISU or DO relative clauses. Relative clauses that have the highest frequency (100) are OBL-ISU, as is shown in Table 19.

Table 19: Positions of Heads in Relative Clauses and in Main Clauses

	TSU	ISU	DO	IO	OBL	GEN	OCOMP
SU(TOPIC)	16	26	20	0	16	4	1
SU	29	40	43	0	20	20	4
DO	46	85	98	0	44	14	2
OBL	44	100	98	0	98	19	0
GEN	19	48	46	1	32	10	0
OCOM	2	3	3	0	2	0	0
SUCOM	13	17	37	0	17	3	1
OTHERS	5	2	5	0	7	6	0

Thus, the results of CA show some agreement with the overall results of this study in that, first, the percentage of definite relative clauses is higher than indefinite relative clauses; second, relative clauses that refer to inanimate referents are more frequent than relative clauses referring to animate referents. On the other hand, the distribution of relative clauses in CA is different from the one reflected by the overall data of this study.

5.6.2 MSA

The MSA texts yielded 1451 relative clauses, which make up to 52% of all relative clauses found in Arabic texts. While the differences between ISU, TSU and DO are not significant, the differences are significant between ISU and OBL (*Wald's z*=-2.155, p=0.031), ISU and GEN (*Wald's z*=-5.142, p<0.001) and between OBL and GEN (*Wald's z*=-5.378, p<0.001*)* (see Tables 56 and 57 in Appendix B). As is shown in Figure 16, the MSA line (in red) shows that the DO position appears at the highest point which indicates that it has the highest frequency. There are slight differences between ISU, TSU and DO. The OBL and GEN positions are significantly lower than the three upper positions. The difference between ABS and SU relative clauses in MSA is not significant (see Table 58 in Appendix B).

Definite relative clauses in the MSA texts (84.35%) outrank indefinite relative clauses, and the results in Table 55 in Appendix B show that the difference between the two categories is highly significant (*Wald's z*=23.274, *p*<0.001). The distribution of definite relative clauses is different from the distribution of indefinite relative clauses. GEN relative clauses are used less frequently than the other relative clauses in the distribution of definite relative clauses while there is no significant difference among the frequencies of indefinite relative clauses, (see Tables 59-62 in Appendix B).

Relative clauses that modify inanimate referents (80%) are more frequent than relative clauses that modify animate referents (19%), and the difference between the two categories is found to be significant (*Wald's z*= 21.279, *p<0.001*). On the other hand, the frequency of relative clauses modifying nonhuman animate referents is only 1%. The distribution of relative clauses that identify animate referents is different from the distribution of relative clauses that identify inanimate referents. Whereas TSU and ISU are used significantly more frequently than DO and other relative clauses in the distribution of relative clauses whose heads refer to animate referents, DO are used more frequently than TSU and other relative clauses refer to inanimate referents (see Tables 63-66 in Appendix B).

Nonfiction relative clauses (54%) outrank fiction relative clauses (46%) in MSA texts. Similar to CA, the MSA data suggest that the difference between fiction and nonfiction is significant (*Wald's z*=2.364, *p*=0.018; see Table 56 in Appendix B). The distribution of fiction relative clauses is similar to the overall distribution of relative clauses in MSA texts in that GEN are less than OBL relative clauses, which are used significantly less than other relative clauses. On the other hand, in the distribution of nonfiction relative clauses, there is no difference between ISU and OBL relative

clauses, which are used less frequently than DO and TSU (see Tables 68- 71 in Appendix B), as is illustrated in Figure 18.

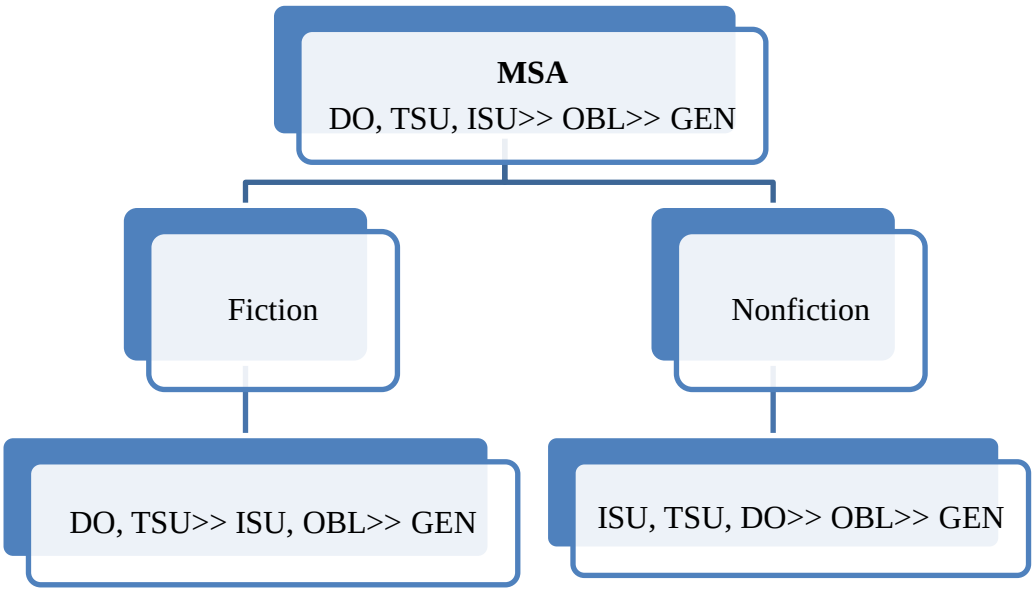


Figure 18: The distribution of MSA relative clauses in fiction and nonfiction texts

Hardy and Milton (1994) suggest that the distribution of relative clauses in first-person narrative is worth studying. The authors state that DO relative clauses are predicted to have the highest frequency due to the presence of exophoric reference in these texts. The MSA fiction texts in this study are written in first-person narrative. As is shown in Figure 25, the frequency of DO relative clauses is significantly higher than that of ISU relative clauses (*Wald's z*=-2.943, *p*=0.003), and also higher than that of TSU relative clauses, yet the difference between the two is not significant (See Table 68 in Appendix B). For a discussion of how Hardy and Milton's (1994) prediction is borne out by the results of MSA fiction, refer to section 6.5.

The data from the MSA texts show that most of the heads of relative clauses (33%) occupy the OBL position in the main clauses. The OBL position is followed,

respectively, by GEN, DO, SU(TOPIC), SU, SUCOM and OTHERS which includes (EXIST, IO, OCOM, OCOMP, CACC), as is shown in Figure 19.

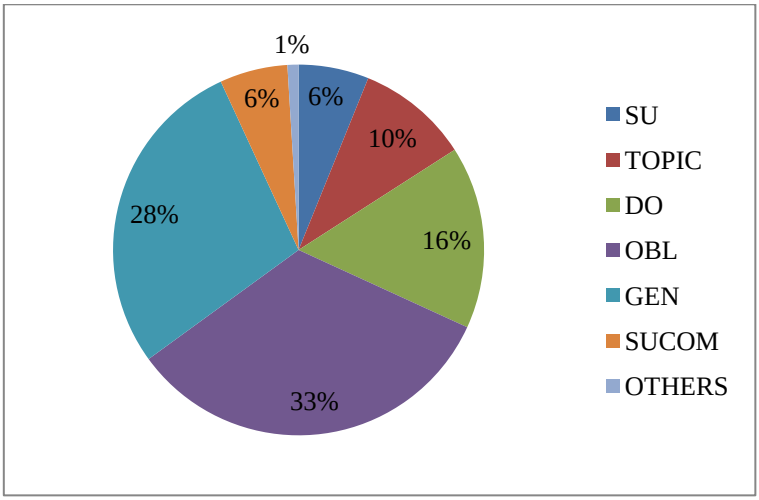


Figure 19: The classification of CA relative clauses according to the grammatical positions of their heads

Similar to CA, the results show that the difference between OBL heads in MSA and heads in other grammatical positions is significant (see Table 72 in Appendix B). The heads of relative clauses show different tendencies to occupy certain grammatical positions in the relative clauses, as is shown in Table 20.

Table 20: Grammatical Positions of the Heads in the Relative Clauses and in the Main Clauses in MSA Texts

	TSU	ISU	DO	IO	OBL	GEN	OCOMP
SU(TOPIC)	45	34	49	1	12	1	0
SU	22	38	19	0	4	6	0
DO	66	49	74	0	26	15	0
OBL	129	116	114	0	105	18	0
GEN	99	106	126	0	51	26	0
SUCOM	24	15	26	0	15	6	0
OTHERS	6	2	3	0	3	0	0

As is shown in Table 20, heads that occupy the grammatical positions with the highest percentages have the tendency to be identified by DO relative clauses except for heads that occupy the SU, OBL and OTHERS positions in the main clause. The OBL heads have a strong tendency to be identified by TSU, then ISU, followed by DO relative clauses. OBL-TSU relative clauses (129) have obtained the highest frequency among other combinations found in MSA relative clauses.

5.6.3 IA

The results of the IA counts yielded 168 relative clauses, which make up only 6.03% of the data of this study. This low percentage is due to two reasons, first, finding books written in Iraqi Arabic was difficult (refer to sections 2.3 and 4.2), and, second, it might be attributed to the method of data collection in IA texts, where only the conversations among characters were counted (see section 4.2).

Relative clauses in IA texts are found in the following descending order, DO, ISU, TSU, OBL, GEN, as is shown in Figure 15. The differences among relative clauses is found significant only between ISU and GEN (*Wald's z*=-2.998, *p*=0.003), as is shown in Tables 73 and 74 in Appendix B. The ABS relative clauses are used more than SU relative causes, yet the difference between these two categories is not significant (see Table 75 in Appendix B).

The percentage of definite relative clauses (70%) outranks the percentage of indefinite relative clauses, and the difference between the two categories is significant (*Wald's z*=4.947, *p<0.001*). The distribution of indefinite relative clauses and the overall distribution of IA do not show any significant differences among the grammatical positions. On the other hand, in the distribution of definite relative clauses,

there is a significant difference between DO and ISU, and the difference between OBL and GEN is approaching significance (see Tables 76- 79 in Appendix B).

In relation to animacy, relative clauses that modify inanimate referents (61%) are found to be the most frequent relative clauses, followed by relative clauses that refer to human referents (35%), and relative clauses that modify nonhuman animate referents (4%) come last. The results show that the difference between animate and inanimate relative clauses is statistically significant: (*Wald's* z=2.755, p<0.006; see Table 73 in Appendix B). DO relative clauses are the most frequent among relative clauses identify inanimate referents. On the other hand, there is no significant difference among TSU, ISU, DO and GEN in the distribution of relative clauses whose heads refer to animate referents (see Tables 80- 83 in Appendix B).

Unlike the other Arabic varieties included in this study, the highest percentage of IA relative clauses occur with SUCOM heads. The OBL position, which has the highest percentages in other varieties, comes in the third place after SUCOM, then SU(TOPIC) and DO, which both occupy second place, as is shown in Figure 20.

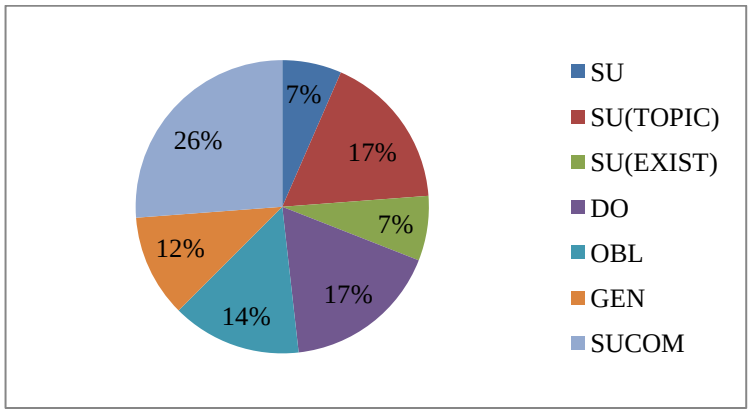


Figure 20: The classification of IA relative clauses according to the grammatical positions of their heads

Statistical analysis of the grammatical positions of the heads of relative clauses show that SUCOM is significantly more frequent than all other grammatical positions (see Table 84 in Appendix B). The most frequent combination found in IA texts is SUCOM-DO relative clauses (17), as is clearly demonstrated in Table 21.

Table 21: Grammatical Positions of Heads in Relative Clauses and in Main Clauses in IA Texts

	TSU	ISU	DO	IO	OBL	GEN	OCOMP
SU	2	4	3	0	2	0	0
SU(TOPIC)	5	6	10	0	6	2	0
DO	8	4	12	0	3	2	0
IO	0	0	0	0	0	0	0
OBL	3	6	9	0	4	2	0
GEN	5	4	6	0	3	1	0
SUCOM	11	7	17	0	5	4	0
OTHERS	4	6	0	0	0	2	0

5.6.4 Summary

The previous sections show the results of the use of relative clauses in the three varieties of Arabic. The results suggest that the order of relative clauses differs across varieties, yet the AH is reflected in the distribution of the three varieties, as is shown in Figure 21 (for a discussion of the results from the different varieties of Arabic, refer to section 6.4. As is shown in Figure 21, none of the varieties matches the overall distribution of relative clauses. MSA appears to be similar to IA except that there is a significant difference between GEN and OBL and the other grammatical positions in MSA, but not in IA. On the other hand, CA is different from the other two varieties as the TSU position is not in the three top positions, while the OBL position is, and the difference between OBL and TSU is significant.

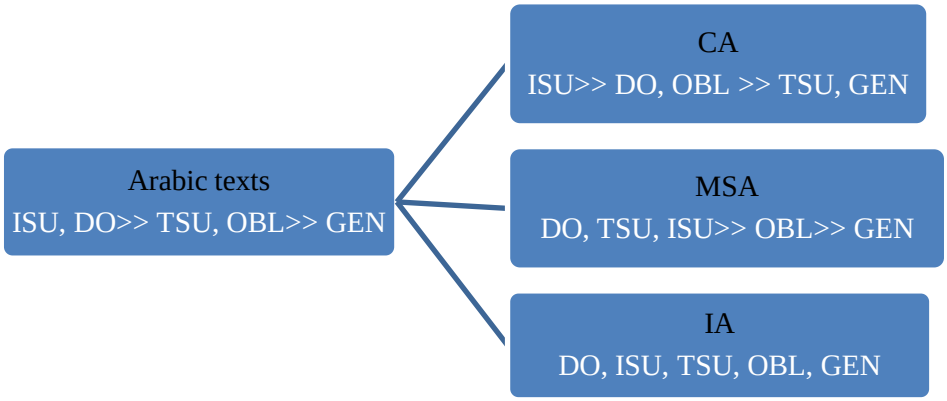


Figure 21: The distribution of relative clauses in the three varieties of Arabic

The varieties of Arabic share some points: First, definite relative clauses significantly outrank indefinite relative clauses in the three varieties; second, the percentage of relative clauses which identify inanimate referents significantly outrank the percentages of relative clauses which identify (human and nonhuman) animate referents in the three varieties of Arabic. No pattern can be found in the distributions of definite and indefinite relative clauses, yet there is a kind of consistency in the distribution of animate and inanimate relative clauses across varieties, as is shown in Figure 22, below.

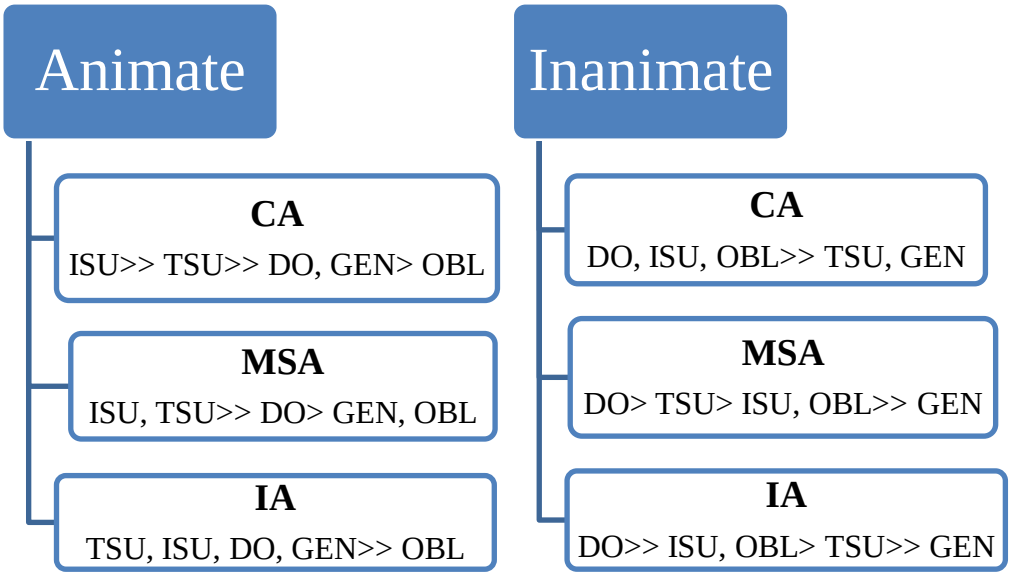


Figure 22: The distribution of relative clauses that refer to animate and inanimate referents in the three varieties of Arabic

As is shown in Figure 22, ISU and TSU, therefore SU, relative clauses are the most frequent among relative clauses that identify animate heads. Furthermore, GEN relative clauses are used more frequently than OBL relative clauses in two of the varieties, CA, IA. On the other hand, DO relative clauses are the most frequent in inanimate relative clauses in two varieties, MSA and IA, and the OBL position is found to be higher than the TSU position in two varieties, CA and IA.

Regarding the position of the head in the main clause, the results show that relative clauses in CA, as well as MSA, are preferred in OBL in the main clause, while relative clauses of IA are preferred in SUCOM position; the OBL position in IA comes fourth. Relative clauses with heads occupying OBL position are modified mostly by ISU relative clauses in CA, TSU relative clauses in MSA, and DO relative clauses in IA.

5.7 Genre

There are two genres included in this study, fiction and nonfiction. First, I will examine, in this section, the impact of genre across the data of this study. Then in each variety, I examine the distribution of relative clauses and whether that distribution reflects the NPAH or the AH. I also examine other factors: the definiteness of the head, the position of the head in the main clause and the animacy of the head, and how these factors are interrelated with the two kinds of genre included in this study.

As is shown in Model 1, the difference between the two genres is found to be significant (*Wald's z*=4.281, $p<0.001$). The interaction between relative clauses and genre will be tested in Model 8, as is shown in table 22.

Table 22: Model 8: Mixed Effect Poisson Regression for the Interaction between Relative Clauses and Genres

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.266	0.144	8.737	<0.001*
TSU	-0.065	0.215	-0.304	0.761
DO	0.202	0.150	1.351	0.177
OBL	-0.298	0.139	-2.147	0.032 *
GEN	-1.342	0.202	-6.656	0.011*
NONFIC	0.520	0.173	3.016	0.002 *
IA	-1.310	0.225	-5.833	<0.001*
MSA	-0.253	0.128	-1.967	0.049 *
DEF	0.987	0.043	23.105	< 0.001*
INANIM	1.060	0.043	24.379	< 0.001*
TSU:NONFIC	-0.692	0.332	-2.088	0.037 *
DO:NONFIC	-0.547	0.226	-2.417	0.016 *
OBL:NONFIC	-0.543	0.206	-2.634	0.009 *
GEN:NONFIC	-0.534	0.303	-1.764	0.078 .

There are five predictors tested in Model 8: (1) relative clauses, ISU, TSU, DO, OBL, GEN, (2) genre, (3) variety, (4) definiteness and (5) animacy. As is shown in rows 2-5 of Table 21, the difference between ISU relative clauses and other relative clauses is not significant except OBL (*Wald's z*=-2.147, p=0.032) and GEN (*Wald's z*=-6.656, p=0.011). The difference between the two genres is also significant (*Wald's*

z=3.115, *p<0.001*). Similar to the previous models, the results in Model 8 show that the differences between definite and indefinite, animate and inanimate, CA and the other two varieties are significant.

The last four rows of Table 22, which show the interaction between relative clauses and genre, indicate that there are some significant interactions between TSU, DO, OBL and GEN on one hand and nonfiction on the other. These interactions are shown in Figure 23.

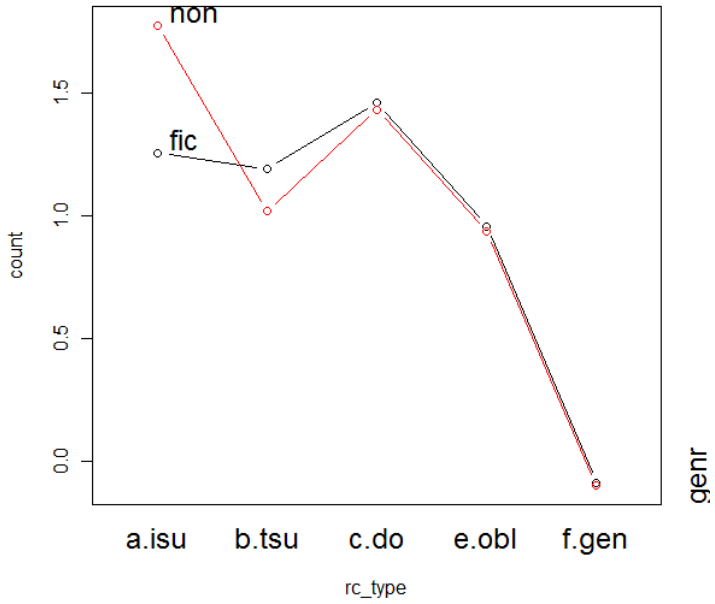


Figure 23: The interaction between relative clauses and genre in Arabic texts, where 'fic' represents fiction and 'non' represents nonfiction

Figure 23 has five points which represent ISU, TSU, DO, OBL, GEN. There are two lines: black, which represents the distribution of fiction relative clauses, and red, which represents the distribution of relative clauses in nonfiction texts. As is shown in Figure 23, the two lines appear identical in the distribution of DO, OBL and GEN, which indicates that fiction and nonfiction relative clauses differ in the frequencies of

ISU and TSU relative clauses. The ISU position is quite high in comparison with other positions in fiction and nonfiction lines.

In the following sections, the results of the distribution of relative clauses in fiction and nonfiction texts are presented. The results of definiteness, animacy and the grammatical position of the heads are also analysed and presented in each genre.

5.7.1 Fiction

The fiction texts yielded 1393 relative clauses, which makes up to 50% of all relative clauses in the sample. The results of the fiction texts yield the following descending order of frequency: DO, ISU, TSU, OBL, GEN, as is shown in Table 85 in Appendix B.

DO relative clauses have the highest frequency among other relative clauses in fiction texts. The difference between DO, ISU and TSU is not significant, and the difference between the TSU position and the OBL position is approaching significance (*Wald's* z=1.708, p=0.088; see Table 86 in Appendix B).

Testing the difference between ABS relative clauses and other relative clauses (see Table 87 in Appendix B) reveals that ABS relative clauses come first before SU relative clauses and the difference between the two categories is significant (*Wald's* z=2.140, p=0.032). Definite relative clauses (1010/73%) outrank the number of indefinite relative clauses (383) in fiction texts, and the difference between the two is highly significant (*Wald's* z=16.167, p< 0.001). Definite DO relative clauses are used more frequently than other relative clauses. On the other hand, there is no significant differences among different indefinite relative clauses (see Tables 88-91 in Appendix B).

The results also show that the majority of relative clauses in fiction texts modify inanimate referents (76%), followed by relative clauses which modify human referents (21%), and then relative clauses that modify nonhuman animate referents (3%). The results presented in Table 85 in Appendix B suggest that the difference between relative clauses that identify animate referents and those that identify inanimate referents is significant (*Wald's z*= 18.525, *p*< 0.001). The distribution of relative clauses referring to inanimate referents shows that DO relative clauses are used more frequently than other relative clauses. On the other hand, there is no significant difference among TSU, ISU and DO in the distribution of animate relative clauses in fiction texts (see Tables 92-95 in Appendix B).

Data from the fiction texts show that the heads of most relative clauses occupy the position of OBL followed by GEN and DO then SUCOM, SU(TOPIC), SU, and OTHERS, respectively, as is shown in Figure 24. The differences between the OBL position and other grammatical positions of the heads are statistically significant (see Table 96 in Appendix B).

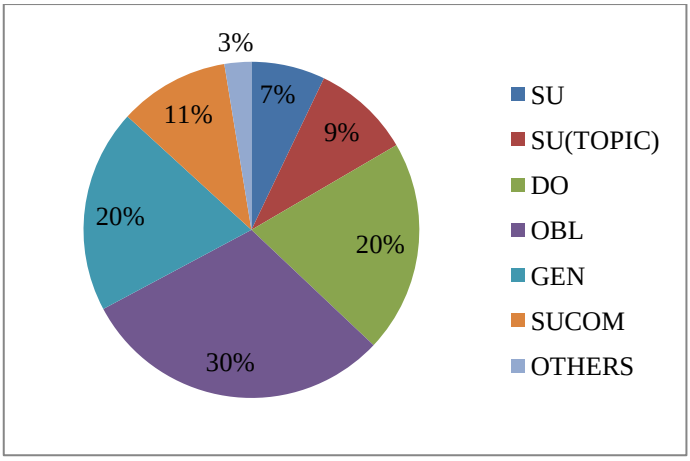


Figure 24: The classification of relative clauses in fiction texts according to the grammatical positions of their heads in the main clauses

Preferences to be identified by certain relative clauses are different among noun phrases in different grammatical positions. Noun phrases in the OBL are modified mostly using ISU relative clauses. The most frequent combination found in the data of this study is OBL-ISU relative clauses (106), as is displayed in Table 23.

Table 23: Grammatical Positions of Heads in the Relative Clause and in the Main Clause in Fiction Texts

	TSU	ISU	DO	IO	OBL	GEN	OCOMP
SU	28	26	25	0	14	6	0
SU(TOPIC)	26	36	41	1	25	3	0
DO	73	61	85	0	43	21	2
OBL	93	106	100	0	99	22	0
GEN	55	64	82	0	54	18	0
SUCOM	33	26	53	0	25	10	1
OTHERS	13	10	5	0	4	3	1

5.7.2 Nonfiction

Nonfiction texts yielded 1392 relative clauses, which appear in the following descending order: ISU, DO, TSU, OBL, GEN. The differences between ISU relative clauses and other relative clauses are significant in all of the grammatical positions tested, as is shown in Table 97 in Appendix B. The results show that ABS relative clauses are more frequent than SU and other relative clauses and the difference between ABS and SU relative clauses is approaching significance (*Wald's* z=1.673, p=0.094; see Table 100 in Appendix B).

The nonfiction texts yielded 1012 definite relative clauses, which outnumber the indefinite relative clauses (380), and the difference between the number of definite and indefinite relative clauses is significant (*Wald's* z=16.503, p< 0.001; see Table 97 in

Appendix B). The distribution of definite relative clauses shows that there is no significant difference between ISU and DO relative clauses, which both are used more frequently than other relative clauses. On the other hand, ISU relative clauses are the most frequent relative clauses in the distribution of indefinite relative clauses in nonfiction Arabic texts (see Tables 101-104 in Appendix B).

The results of the nonfiction texts show that the majority of relative clauses (72%) modify inanimate referents followed by relative clauses whose heads refer to human referents (26%), and few relative clauses (2%) refer to nonhuman animate referents. The difference between relative clauses that identify animate referents and those which identify inanimate referents is highly significant (*Wald's* z=15.893, $p<$ 0.001), as is shown in (Table 97 in Appendix B).The distribution of relative clauses that refer to animate referents shows that ISU relative clauses are the most frequent followed by TSU and then DO relative clauses, which come in the third position. In the distribution of relative clauses referring to inanimate referents, DO and ISU relative clauses come first (see Tables 105-108 in Appendix B).

Similar to the fiction texts, most of the heads of relative clauses in nonfiction texts occupy the OBL position in the main clause, then, come relative clauses whose heads occupy the GEN position, as is shown in Figure 25. The differences between the OBL position and other grammatical positions are significant, as is shown in Table 109 in Appendix B.

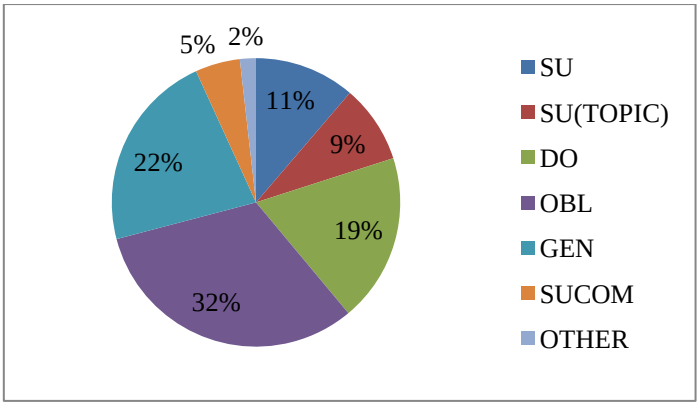


Figure 25: The classification of relative clauses in nonfiction texts according to the positions of their heads.

Noun phrases that occupy the OBL grammatical position in the main clause have the tendency to be identified by ISU relative clauses. The most frequent combination is found in OBL-ISU relative clauses (130), as is shown in Table 24.

Table 24: Grammatical Positions of Heads in the Relative Clause and in the Main Clause in Nonfiction Texts

	TSU	ISU	DO	IO	OBL	GEN	OCOMP
SU	25	74	22	0	12	20	4
SU(TOPIC)	40	31	37	0	9	4	1
DO	46	104	73	0	30	10	0
OBL	83	130	107	0	108	17	0
GEN	67	100	91	1	32	19	0
SUCOM	15	17	23	0	12	3	0
OTHERS	4	4	5	0	7	4	0

5.7.3 Summary

Although the AH is reflected in the distribution of relative clauses in fiction and nonfiction texts, there are slight differences between these two genres. Whereas the ISU has the highest frequency in nonfiction texts, the DO relative clauses are the most frequent relative clauses in fiction texts, although the difference between the DO, ISU and TSU is not statistically significant, as is shown in Figure 26 (for a discussion of the results of genre refer to section 6.5).

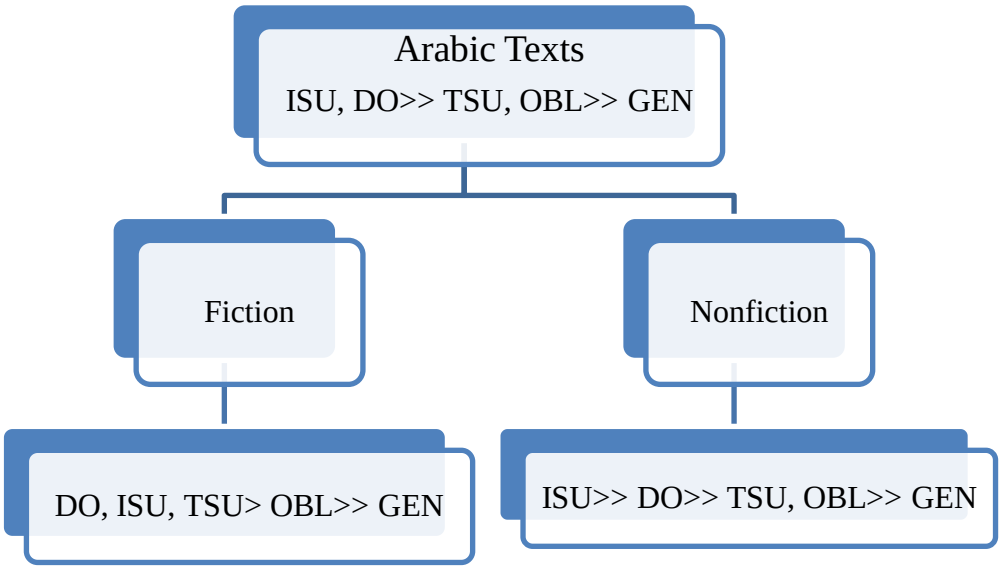


Figure 26: The distribution of relative clauses in fiction and nonfiction texts

Definite relative clauses outrank indefinite relative clauses in both genres; this indicates that the tendency towards including more definite relative clauses than indefinite relative clauses is similar in Arabic fiction and nonfiction texts.

Similarly, relative clauses that modify inanimate referents outrank other relative clauses in the two genres with similar percentages (76%) and (72%) in fiction and nonfiction, respectively. Thus, the tendency to identify inanimate referents is similarly found in both genres. Furthermore, fiction and nonfiction texts appear to have similar tendencies as far as the positions of the head in the main clause. In both genres, the heads of relative clauses occupy the OBL position in first place with similar percentages, (30%), and (32%) in fiction and nonfiction, respectively. In the two genres, the OBL heads tend to be identified by ISU relative clauses.

5.8 Summary

The results show that the combined values of TSU and ISU are significantly higher than relative clauses in other grammatical positions, DO, OBL, GEN, yet the AH is supported more than the NPAH. ABS relative clauses are used more frequently than SU and other relative clauses in Arabic written texts. However, this has not been consistent over the three varieties since the difference between ABS and SU is statistically significant in CA, but not in MSA and IA.

The results of this study show that that there is no consistency in the order of frequency of relative clauses among varieties and between genres. The results of varieties show that relative clauses in each variety show an order of frequency, which is different from the orders shown by the other varieties. In particular, CA is different from the other two varieties since OBL relative clauses are used significantly more frequent than TSU and GEN. Similarly, the order of relative clauses in fiction is different from the order in nonfiction.

The results of the definiteness and animacy counts suggest that these two factors contribute to the distribution of relative clauses in this study. Definite relative clauses are found significantly more frequently than indefinite relative clauses, and relative clauses in each category show a different frequency order of relative clauses. Similarly, relative clauses that identify inanimate referents occur significantly more frequently than relative clauses that identify human and nonhuman animate referents.

Animacy seems to play a more important role, first, in deciding the distribution of relative clauses and, second, in deciding whether the order of relative clauses follows the subject primacy hypothesis, the NPAH or the AH. Relative clauses referring to animate referents tend to have SU (ISU or TSU) relative clauses in the highest position,

whereas the DO position mostly comes first in the order of frequency of relative clauses referring to inanimate referents. GEN relative clauses are more frequent than OBL in the order of relative clauses referring to animate referents in fiction texts and in two varieties, CA and MSA, which violates the order of the NPAH.

The results for the grammatical position of the head show that heads of most relative clauses occupy the OBL position. OBL heads mostly have the preference to be identified by ISU relative clauses or DO relative clauses. These results show a kind of consistency among genres and varieties except IA, where heads of relative clauses tend to occupy the SUCOM position.

Chapter Six: Discussion

6.1 Introduction

One of the main goals of this study is to analyse the distribution of relative clauses in Arabic written texts. In particular, this study aims at checking whether the variety of Arabic and genre have any effects on that distribution. It also tests the role of other factors, which we called construction-internal factors, in defining the shape of distribution of relative clauses. The analysis of the results is presented in Chapter Five, and this chapter will discuss the findings in light of the aims of this study and in relation to previous studies. First, a general discussion of the results is presented in section 6.2. Second, whether these results follow the NPAH or the AH is discussed in 6.3. Then, the effect of construction-external factors, varieties and genres, on the distribution of relative clauses is discussed in sections 6.4 and 6.5, respectively. Next, the influence of construction-internal factors on the use of relative clauses in written texts is analysed in section 6.6. Finally, a summary of the discussion chapter is presented in section 6.7.

6.2 General Results and Counts

As was shown in section 5.2, relative clauses in Arabic texts show the following descending order: ISU, DO, TSU, OBL, GEN, OCOMP, IO. In fact, the IO position is the least frequent position; this is counter to the NPAH's order, which is as follows: SU >DO> IO> OBL> GEN> OCOMP.

Hogbin and Song (2007) revealed similar results where no instances of IO relative clauses were reported. The authors offered two possible reasons for that deviation, which I will examine in light of the data of this study. First, the indirect

object is used infrequently in main clauses in discourse; therefore, it might be expected that IO relative clauses would occur infrequently if at all. Second, indirect objects tend to be pronominal and relativisation on the grammatical relations that tend to be pronominal, such as TSU, IO and GEN, is not frequent.

There is a tendency for subjects to be pronominal, yet the SU position has the second highest percentage of relative clauses, hence the first reason is applicable to this study whereas the second reason is not. The results show that the number of relative clauses which have IO heads is only 6 out of 2785. This can be attributed to the fact that indirect object is restricted to the role of beneficiary or recipient and it is also connected with human or animate referents (Palmer, 1994).

Relating the IO position to animacy might give another explanation for the infrequent use of IO relative clauses in Arabic discourse; this can be explained through the following two reasons:

1. The percentage of heads that refer to inanimate referents is found to be significantly higher than the percentage of animate referents (see section 5.5).
2. When identifying human beings, there is a strong preference to refer to their actions or their characteristics (i.e. TSU and ISU) rather than identifying them by referring to the benefit that they have obtained (i.e. IO relative clause) unless a writer wants to focus on the benefit rather than the characteristics.

Thus, in light of the small number of relative clauses referring to animate, and IO's restriction to beneficiary role, IO relative clauses are not used frequently in discourse.

The strong preference toward using SU rather than IO relative clauses might be attributed to the fact that the SU is included in the structural domain or the relativisation domain of IO, but the IO is not included in the relativisation domain of SU. That is, the IO relative clause has to include the SU whereas it is not necessary for the SU relative clause to occur with the IO (Hawkins, 1994). In the IO relative clause, the writer needs to identify the subject and direct object, before identifying the IO. In contrast, in SU relative clauses, the writer does not need to refer to the DO or to the IO. Thus, SU relative clauses are easier to produce and process than IO relative clauses.

OBL relative clauses, on the other hand, have a higher percentage of distribution in the corpus in spite of the peripheral nature of the OBL position. This can be explained in a parallel way to the explanation for the infrequent occurrence of IO relative clauses, but in the opposite direction. Namely, OBL relative clauses are related to time, place, and means among other things; this means the preference of the OBL position is to have inanimate heads. OBL relative clauses are used in discourse to describe the scene (Hogbin & Song, 2007, p. 222). For example, when writers want to identify a certain place, time or means they mainly use OBL relative clauses, as in examples (85) and (86), which are OBL relative clauses of place and time, respectively.

85.	...wa	al-mawḏiʿ	allaḏī	wulida	fī-hi
	...and	the-place	REL(3.SG.MAS)	born	in-it

...and the place in which he was born.

[CA, 4: 81]

86. wa mata fī al-yawmi allaḏī mata fī-
and died in the-day REL(3.SG.MAS) died in-
-hi Abu Bakr
-it Abu Bakr
And he died in the day in which Abu Bakr died.
[CA, 5: 305]

The results of this study show that OBL relative clauses that refer to place have the highest percentage (39.02%), followed by OBL relative clauses that refer to time (22.36%), which confirms the fact that one of the important functions of OBL relative clauses is to set the scene. Since, setting a scene is equally important in both novels and history books, OBL relative clauses are found with high frequency.

GEN relative clauses are not used frequently in Arabic texts as they have a percentage of 5.78% only. This percentage agrees with the NPAH's predictions since the GEN position is the second lowest position on the NPAH. However, using GEN relative clauses more frequently than IO relative clauses is against the NPAH's predictions.

One possible answer to the high frequency of GEN relative clauses in comparison to IO relative clauses is that generally the GEN structure, outside relative clauses, is considered as "one of the most basic structures in Arabic and occurs with high frequency" (Ryding, 2005, p. 205). A comparison between GEN heads (i.e. heads of relative clauses that occupy the GEN position in the main clauses) and IO heads in the data of this study shows that while GEN heads are found with a percentage of 21%, IO heads are found with a percentage of 0.02% (see section 5.4). This suggests that GEN noun phrases in general are found more frequently in written texts than the IO noun phrases.

It might be argued that what has led to this result is including more types in this study than Keenan and Comrie (1977) included in their study, which is the possessive relation only (see section 4.4.1.5). Considering the percentages of the kinds of GEN included in this study, I have found that including possessors only and overlooking other kinds of GEN, which make up to 50% of GEN relative clauses, will not result in a representative sample of GEN relative clauses, although the possession relationship has the highest percentage among other kinds of relationships. In fact, including only the possession relationship will reduce the percentage of the GEN position to 2.78% instead of 5.78%, and the difference between the two percentages is statistically significant. However, this difference does not affect the distribution of relative clauses found in Arabic written texts if compared to the NPAH. The GEN position (whether relationships other than possession are included or not) follows the OBL and comes before the OCOMP and the IO. The violation in the IO position would still exist if only the possession relation was included, as in both cases, GEN relative clauses occur more frequently than IO relative clauses.

Finally, there are only eight OCOMP relative clauses in the Arabic texts, which makes up 0.28% of the sample. Based on the NPAH's predictions, this percentage is expected. All of the OCOMP relative clauses found are used to express the superlative in a way which is combined with a little exaggeration by using the negative 'never seen ...better', as is seen in example (87).

87. qad nazal bi-nā qaum lam narā
already came down in-us people not we.see
qawm qaṭ aḥsan min-hum
people never better than-them
We have never seen better than the people who came to our house.
[CA. 4: 103]

This style of expressing the superlative through OCOMP is found only in CA rather but not in MSA and IA, which, first, highlights the stylistic changes among the varieties (Ryding, 2005), and also supports the claim that CA uses more complex structures than MSA (Bassiouney, 2009a).

6.3 The NPAH or the AH

This section discusses the results in relation to the NPAH and the AH. It answers two questions: (1) Does the distribution of relative clauses in Arabic texts follow the NPAH? and (2) Does the distribution of relative clauses in Arabic texts follow the AH? These questions are addressed in 6.3.1 and 6.3.2, respectively.

6.3.1 The Results and the NPAH

When we look at the entire corpus, there is some support for the NPAH. Merging TSU and ISU into one position, the SU position, as is stated in the NPAH, supports the NPAH since the percentage of relative clauses formed on the SU position is found to be significantly higher than the percentages of relative clauses formed on the other grammatical positions on the NPAH (see section 5.2.1); this result supports the SPH. This result also conforms with a number of previous studies which state that SU relative clauses are easier to process than DO relative clauses (e.g. Kuo and Vasishth (2010); Kwon et al (2010) on Korean relative clauses; Lin and Bever (2006) on Chinese relative clauses; Ishizuka (2005) and Ueno and Garnsey (2008) on Japanese relative clauses).

The results of this study show a similarity to some previous corpus studies. Among these studies are Keenan (1975), Gordon and Hendrick (2005) and Hogbin and

Song (2007), where TSU and ISU are merged into one position and the SU position has the highest percentage of frequency among the grammatical positions on the NPAH. All of these studies suggest that SU relative clauses are used more frequently than other relative clauses in English written corpora.

In this study, the results of other positions down the NPAH suggest that there is a gap shown in the IO position since it occurs only twice, although it is the third position in the NPAH (see sections 5.2 and 5.2.1). Hogbin and Song (2007) obtained similar results where there was no occurrence of IO relative clauses in their sample of English texts. However, the results in both Keenan (1975) and Gordon and Hendrick (2005) have not shown any gap in the NPAH.

One reason for the difference between the results of this study and Keenan's results is that Keenan (1975) included IO relative clauses under the OBL position because the author believed that, in English, IO and OBL positions behave similarly with respect to relativisation. In the same way, Gordon and Hendrick (2005) tested three positions only, the SU, DO, and OBL positions so there is not a separate position for IO. However, IO and OBL relative clauses are tested in separate positions in Hogbin and Song (2007) where the authors state that " there is little reason, in written English, to collapse the two positions, not least because there is so substantial a difference in frequency between them" (2007, pp. 227-228). In Arabic, IO relative clauses can be distinguished from OBL relative clauses, so in this study IO and OBL are placed into two different categories (see sections 4.4.1.3 and 4.4.1.4). This implies that it is possible that if Keenan (1975) and Gordon and Hendrick (2005) had tested IO relative clauses as a separate position, they would have reached similar results.

The fact that the IO position is found out of the hierarchical order makes it possible to argue against the NPAH. However, the results of this study alone cannot

invalidate the NPAH because the distribution of relative clauses in discourse might not be related to accessibility to relativisation only. Among the factors that could contribute to the frequency of relative clauses in the text is the function that relative clauses have, which will be discussed in section 6.4. Nonetheless, there is stronger evidence for the AH, since the SPH is not supported in certain parts of the data, as will be discussed next.

6.3.2 The Results and the AH

The overall results of this study indicate that ABS relative clauses (ISU+DO) are used more frequently than other relative clauses on the NPAH (TSU, IO, OBL, GEN, and OCOMP) (see section 5.2.2). SU relative clauses (TSU+ ISU) (49.91%) are found less frequently than ABS relative clauses (ISU+ DO) (55.22%), and the difference between the two types of relative clauses is significant (see sections 5.2.1 and 5.2.2). This result is found consistently in the two types of genres, fiction and nonfiction, but not in the all of the three varieties of Arabic. ABS relative clauses are used significantly more than other relative clauses including SU in CA, and not in MSA and IA where the difference between ABS and SU is not significant. Thus, there are some results that support the AH, and other results that neither refute nor confirm the AH; these results confirm the effect of varieties on the distribution of relative clauses in the texts (as will be discussed in the next section).

The results of the distribution of relative clauses in the corpus as a whole gives some support to Fox's assertion that "it seems to be the category ABSOLUTIVE, rather than SUBJECT, which occupies the leftmost position on the accessibility hierarchy" (Fox, 1987, p. 869). This has been confirmed by Hogbin and Song, who state that "Perhaps of most significance is the fact that S-RCs (ISU relative clauses) and DO-

RCs—with no statistical difference between the two—occurred more frequently in both the eighteenth- and twentieth-century sample texts than did A-RCs (TSU relative clauses)" (Hogbin & Song, 2007, p. 229). Gordon and Hendrick (2005) also found that the AH is manifested in two out of the three corpora which are included in their study. Thus, the AH has support not only from spoken English discourse but also from written discourse in English and Arabic.

Fox (1987) explains that relative clauses help to relate the referent (the head NP) to the whole context through describing or anchoring. While the function that ISU relative clauses have is to describe the referent, both DO and TSU relative clauses have an anchoring function, that is they link or anchor the head into the discourse through another entity in the relative clause, as is shown in examples (88), (89) and (90), which are ISU, TSU and DO relative clauses, respectively.

88. ṣaʿada al-raǧul allaḏī kān ʾilā ǧānib-hā
went up the-man REL(3.SG.MAS) was to side-her
The man who was beside her went up.
[MSA, 1: 34]

89. wa istadʿai-tu al-mallāḥ allaḏī aḫraǧa-nā
and summoned-I the-sailor REL(3.SG.MAS) take out-us
tillka al-laila
that the-night
And I summoned the sailor who took us out that night.
[CA, 1:81]

90. rafaḍ siǧāra-tī allatti qaddam-tu-hā ʾilai-hi
refused cigarette-my REL(3.SG.FEM) presented-I-it to-him

He refused my cigarette that I have presented to him.
[MSA, 2: 10]

Although both TSU and DO relative clauses perform an anchoring function, they fulfil that in two different ways. TSU relative clauses link the referent to the text using the DO inside the relative clause, as in example (89), the head *al-mallāḥ* 'the-sailor' is linked through the object pronoun *nā* 'us'. On the other hand, the head of DO relative clause is linked into the discourse through the TSU inside the relative clause, as in example (90), *siğāra-tī* 'my cigarette' is linked through the subject pronoun *–tu* 'I'.

Fox suggests:

> "P-relatives [DO relative clauses] will by far predominate over A-relatives [TSU relative clauses], because the other NP in a P-relative (i.e. the A [TSU]) is a better anchor than is the other NP in an A-relative (i.e. the P [DO])" (1987, p. 862).

She explains that the anchoring function is better performed by pronouns. In this study, a comparison between DO and TSU relative clauses reveals that DO relative clauses use, as anchors, a significantly higher percentage of pronouns and a lesser percentage of full noun phrases than TSU relative clauses, although both TSU and DO relative clauses have a high tendency to use full noun phrases as anchors (see section 5.2.2). This suggests that DO relative clauses are more suitable to the anchoring function than TSU relative clauses. The results show that, DO relative clauses are significantly more frequent than TSU relative clauses. Thus the results of this study support Fox's (1987) argument:

> "it is the discourse function of S- [ISU] and P[DO]-relatives, rather than a special cognitive status, which gives them prominence in languages across the world" (p. 869).

Song (2001) raised the question of whether the above explanation can be extended into the lower positions of the NPAH (p. 241). Hogbin and Song (2007) suggest that Fox's (1987) claims about the distribution of given and new information can be extended to grammatical positions further down the NPAH. The authors indicate that the other noun phrases in OBL relative clauses tend to be pronominal, and therefore good anchors.

An investigation of the anchoring function in OBL relative clauses has been carried out in this study. Similar to TSU and DO relative clauses, the other noun phrase in OBL relative clauses can serve to anchor the head into the context, as is shown in example (91).

91. .. min	ʾaina	ǧāʾa-t	tilka	al-makānah	allatī
...from	where	came-FEM	that	the-status	REL(3.SG.FEM)
waḍaʿ-ū-hu		fī-ha			
put-they-him		in-it			

From where he got that status in which they put him?
[MSA, 5: 115]

OBL relative clauses show a similarity to DO relative clauses in the way the anchoring function is performed. The referent of the OBL relative clauses can be anchored to the context using the SU of relative clause. As is shown in example (91), the head *al-makānah* 'the-status' is anchored by the subject pronoun *–ū* 'they'. Thus, it can be hypothesised that SUs in OBL relative clauses are better anchors than DOs in TSU relative clauses. The results confirm this hypothesis since the percentage of NPs that are used as anchors in OBL relative clauses is less than the percentage of NPs used as anchors in TSU relative clauses. However, this result does not match the distribution

of relative clauses found in Arabic texts. As is clear in section 5.2.2, there is no significant difference between the frequency of TSU and OBL relative clauses, although OBL relative clauses provide better anchors than TSU relative clauses. This indicates that the distribution of relative clauses cannot be fully explained based on the functions of these relative clauses.

Although the functions of relative clauses affect the distribution of SU and DO relative clauses in discourse, they do not seem to correlate with the distribution of TSU vs. OBL relative clauses. OBL relative clauses provide two functions: characterisation and anchoring (see section 5.2.2). Furthermore, OBL relative clauses provide better anchors than TSU relative clauses, yet they are not used more frequently than TSU relative clauses in Arabic texts. This suggests that the discourse-based explanation cannot be extended to the lower positions of the NPAH, and the "psychological accessibility" or "cognitive prominence" associated with the NPAH might still provide an explanation to the distribution of relative clauses in written texts.

This section has discussed the distribution of relative clauses in Arabic texts in relation to the NPAH and the AH. It is found that the frequency of ISU, TSU and DO relative clauses supports the AH as ABS relative clauses appear more frequently than SU relative clauses. However, the frequency of other relative clauses such as OBL cannot be explained through their functions in the text. It follows that the functions of relative clauses as well as their accessibility to relativisation together play an important role in the distribution of relative clauses in Arabic written texts. The next sections will discuss whether variety of Arabic and the genre can be considered as important variables that influence the distribution of relative clauses in discourse.

6.4 The Role of Construction-External Factors

This section discusses the results in light of the effect of two construction-external factors, variety and genre. In sections 6.4.1, the focus will be put on the question: Does the variety of Arabic in which the relative clauses are written affect the distribution of relative clauses? Then, in section 6.4.2, the question addressed is: Does the genre in which the relative clauses are written influence the distribution of relative clauses?

6.4.1Variety

The results presented in Chapter Five consistently show IA and MSA to be significantly different from CA. The distribution of relative clauses differs from one variety to another (see section 5.6). Whereas CA relative clauses appear in the following descending order ISU>> DO, OBL>> TSU, GEN, relative clauses used in MSA have the following descending order DO, TSU, ISU>> OBL>> GEN; and IA relative clauses appear in the following descending order DO, ISU, TSU, OBL, GEN.

The NPAH is not reflected in any of the three varieties considered separately, especially in CA where the frequency of OBL relative clauses is significantly higher than TSU relative clauses. The frequency of SU relative clauses has not been found higher than ABS relative clauses in any of the three varieties. Therefore, the SPH is not supported in the three varieties. For this reason and the infrequent use of IO relative clauses in the three varieties, which is considered as a violation to the hierarchical order of the NPAH, the results do not confirm Keenan's (1975) claim that the frequency distribution of relative clauses in texts follows the order of the NPAH.

ABS relative clauses are used significantly more than SU relative clauses only in CA. Thus, the AH is manifested in the results of CA, which confirms Fox's (1987)

claim that it is "the category ABSOLUTIVE, rather than SUBJECT, which occupies the left-most position on the accessibility hierarchy"(p. 869). The difference between CA and the other two varieties might be attributed to the stylistic changes and the linguistic structures that are used in CA but not in MSA and IA. These stylistic changes take place due to the differences between the chronological periods of CA on one hand and the other two varieties on the other. An example of these linguistics structures is the use of *yuqāl li-* 'said to-', which is always found in the passive form, to give the meaning of 'called' as in example (92). The use of this verb in passive contributes to the high frequency of ISU in CA.

92.	kān	fī	masǧid	yuqāl	la-*hu*	masǧid
	was	in	mosque	say(PASS)	to-*it*	mosque
	ṣaliḥ					
	Salih					

He was in a mosque which is called Salih's mosque.
[CA, 4: 80]

CA is the only variety of Arabic that Keenan and Comrie (1977) included in their study to represent the Arabic language (p. 76). However, the results of this study show that neither of the varieties can represent the Arabic language. A comparison of the overall distribution of relative clauses in the data of this study with the distributions of relative clauses in each variety does not reveal similar results. That is, the order of relative clauses in either CA, MSA or IA is different from the overall order of the relative clauses in the data of this study, which is ISU, DO>> TSU, OBL>> GEN. Therefore, CA does not sufficiently represent the Arabic language, which conforms with Fleischer's claim that the standard variety is "fairly unrepresentative if compared

to the overall picture" (2004, p. 236). Thus, the variety in which relative clauses are written might not contribute to whether relative clauses follow the NPAH or the AH, yet it is an important factor that influences the general distribution of relative clauses in the texts.

6.4.1.1 Diglossia

The results show that the differences between the distribution of IA relative clauses and MSA relative clauses are not significant in comparison to the differences between CA and IA or CA and MSA. This suggests that, as far as the use of relative clauses, there is no difference between IA and MSA, and both varieties differ from CA. This can be surprising because in looking at the syntax of MSA and CA, we find that relative clauses are formed in the same way using the same relative markers that are inflected to number, gender and case sometimes (3.3). On the other hand, IA relative clauses use relative markers that are not inflected for gender and number (see section 3.4.1). As suggested above in section 6.4.1, the differences between CA and MSA and IA are stylistic differences which are caused by the differences between the chronological periods of CA and the other two varieties.

The results suggests two points: first, since MSA and IA show similarities as far as using relative clauses, switching from MSA into IA can take place frequently in everyday conversations whereas switching to CA is very restricted. This is expected since CA is only used to recite Qur'an or to refer to older books (Bassiouney, 2006, 2009a). Thus, the use of relative clauses is not related to the similarity between the varieties; it is controlled by the contexts in which relative clauses are used in the diglossic situation. Second, the same results can be predicted for the use of relative

clauses in other dialects such as Saudi, Egyptian and Lebanese. That is, the use of relative clauses in these dialects is similar to the use of relative clause in MSA, but not in CA.

The results bring up a question of whether these varieties are treated as different languages in the minds of speakers. This study cannot directly answer this question. However, comparing the distribution of relative clauses in each variety with the distributions of relative clauses in other languages, which are previous studies revealed, shows that the distribution of relative clauses in CA is similar to that in English. Hogbin and Song's (2007) study reveals the following order: ISU/DO> OBL> TSU> GEN and CA relative clauses appear in the following order: ISU>> DO, OBL>> TSU, GEN. Apart from significance, relative clauses in CA and English have the same order. On the other hand, the orders of relative clauses in MSA and IA are different from that in CA. This comparison suggests that, as far as the use of relative clauses is concerned, CA is closer to English than to other varieties of Arabic. Accordingly, it is possible that, in the mind of Arab people, who are exposed to different varieties of Arabic, CA is treated as different language rather than as a variety of the same language, although this cannot be confirmed in this study. This might explain why CA is difficult to learn and use in Arabic world (Ennaji, 2005)

6.4.2 Genre

The results of the fiction and nonfiction counts show that the distribution of relative clauses follows the AH rather than the SPH. SU relative clauses are used less frequently than ABS relative clauses and the difference between the two positions is statistically significant in fiction, and approaching significance in nonfiction (see

sections 5.6.1 and 5.6.2). These results suggest that the AH is reflected in both genres, fiction and nonfiction, although it is manifested in fiction more than nonfiction.

Previous studies such as Hardy and Milton (1994) and Jensen (1999) suggest that genre affects the distribution of relative clauses in written texts. The results of this study suggest that in spite of the similarities between the two genres, there is a significant difference between fiction and nonfiction. The order of relative clauses in fiction texts is different from that in nonfiction texts. ISU relative clauses are the most frequently occurring relative clauses in nonfiction, whereas there is no significant difference between DO, ISU and TSU relative clauses in fiction texts. It follows that there is a relation between the frequency of relative clauses and the way information is presented in the text.

Hardy and Milton (1994) suggest that highly informative texts tend to use ISU relative clauses. The results of this study suggest a similar tendency, as ISU relative clauses are the most frequent in nonfiction texts, which are highly informative. Hardy and Milton also suggest that DO relative clauses are used most frequently in exophoric texts such as conversation and first-person narrative. DO is significantly higher than other relative clauses in MSA texts which are first-person narrative (see sections 5.6.2 and 5.6.3). However, DO is not found higher than ISU in the IA data, which is based on conversations among the characters (see section 4.2). Thus, the relation between exophoric texts and the high frequency of DO relative clauses is partially confirmed by the data of this study.

6.5 The Role of Construction-Internal Factors

In this section a discussion of how factors other than variety and genre such as definiteness, the grammatical position of the head and animacy might contribute to the way relative clauses are distributed in Arabic text is presented. First, in section 6.5.1, I address the question: Does the definiteness of the head affect the distribution of relative clauses in Arabic texts? In section 6.5.2, whether the grammatical position of the head affects the distribution of relative clauses in Arabic texts is discussed. Section 6.5.3, presents an answer to the question: What is the effect of the animacy of the heads of relative clauses on the order of frequency of these relative clauses in the texts?

6.5.1 Definiteness

The results show that definite relative clauses are used more frequently than indefinite relative clauses (see section 5.3). This result shows a consistency across the three varieties of Arabic and the two genres (5.6 and 5.7).

The use of a high percentage of definite relative clauses might be attributed to their ease of processing, which is accomplished in two ways. First, definite relative clauses use relative markers, while indefinite relative clauses do not. The use of the relative marker makes the relative clauses easier to process since the relative marker signals the beginning of another clause in the sentence which is the relative clause. On the other hand the absence of relative marker results in an ambiguous sentence, which will be harder to process (Prideaux & Baker, 1986, p. 39).

Second, in definite relative clauses, the relative marker helps the reader to easily identify the head of the relative clause since the relative marker is used immediately after the head, and the relative markers agree with heads in number and gender (refer to

section 3.3.1.1). Sheldon (1974) suggests that recognising the noun phrase that the relative clause modifies is one of the difficulties which learners face in the processing of sentences with relative clauses (p. 280). Therefore, the presence of a relative marker is an advantage that definite relative clauses have over indefinite relative clauses. To further explain these two points, compare the definite relative clause in (93) with the indefinite relative clause in (94).

93. wa al-ʿağūz allatī tusamā ḏat Al-dawahi
and the-old lady REL(3.SG.FEM) called ḏat Al-dawahi
ğadat-ī um ab-ī
grandmother-my mother father-my
And the old lady who is called ḏat Al-dawahi is my grandmother on my father's side. [CA, 1: 219]

94. ...wa daḫal-nā madīnah tusamā Al-kūḫ fa-biʿ-nā
...and entered-we city call.PASS Al-kūḫ then-sold-we
wa išterai-nā
and bought-we
...and we entered a city which is called Al-kūḫ then we sold and bought (different things).
[CA, 1: 340]

In (93), the reader knows as soon as he reads the word *allatī* that *al-ʿağūz* 'the old lady' is the head of the relative clause and the coming clause is a relative clause so he will process the sentence immediately on that basis. Whereas in (94), the reader has no sign that leads him to identify the head of the relative clause, which is *madīnah* 'a city', because of the absence of the relative marker. This indicates that presence of relative

markers in definite relative clauses make them easier to process and more accessible to comprehension than indefinite relative clauses.

Keenan and Comrie (1977) included only definite relative clauses in their study. On the other hand, both definite and indefinite relative clauses are included in other studies such as Fox (1987) and Hogbin and Song (2007) because the authors of these studies did not find obvious reasons for limiting their studies to relative clauses with definite heads. Due to the difference between the results obtained by Keenan and Comrie (1977) and other studies on one hand and the difference between the structure of definite relative clauses and the structure of indefinite relative clauses in Arabic (refer to sections 2.5.1.1 and 2.4.4.1.2) on the other, it is found to be important to include both types of relative clauses in the data of this study. It is interesting to find out whether definiteness is an influential factor that might affect the distribution of relative clauses in a particular text, and whether including definite relative clauses only reveals the exact order of the NPAH.

ABS definite relative clauses outnumber SU definite relative clauses but are equal for indefinite relative clauses (see section 5.3). This indicates that definiteness has a slight effect on the distribution of relative clauses, yet this does not lead to the manifestation of the NPAH in either of the two categories, definite and indefinite relative clauses. In fact, including definite relative clauses only, as Keenan and Comrie (1977) did, supports the AH more than the NPAH. This is similar to Fox's (1987, p. 861) results where excluding indefinite relative clauses would result in DO relative clauses occurring more frequently than SU relative clauses.

Another question addressed in this study is whether definiteness has a correlation with the grammatical position of the relativised noun phrase. The results

show that there is a correlation between the grammatical position of the relativised noun phrase and its definiteness. ISU, TSU, DO, OBL heads have a strong tendency to be definite. On the other hand, GEN heads have a slight preference, which is not as strong as that in the other grammatical positions, towards being definite.

The results of this study agree with the results obtained by Gordon and Hendrick (2005) who found that ISU and DO relative clauses have definite heads. However, the definiteness of heads does not seem to follow Du Bois's (1987) Preferred Argument Structure, which states that the ISU and DO positions tend to carry new information so they should have the tendency to be indefinite because noun phrases that carry new information tend to be indefinite (Comrie, 1989; Lyons, 1999; Prince, 1981). The Preferred Argument Structure was supported by a number of previous discourse studies (e.g. Fox, 1987; Fox & Thompson, 1990) where the results show that the head nouns in ISU relative clauses have the tendency to be indefinite .

The difference between the results of this study and previous studies might be partially attributed to two factors. First, there is the nature of the Arabic language. In the Arabic language, ISU relative clauses consist either of sentences whose verbs are intransitive or of sentences that do not have any verbs (i.e. equational sentences), which are similar to English sentences that have the verb 'be' as the main verb (refer to section 3.2.2). Badawi et al. (2004) have suggested that the subjects of equational sentences "are nearly always definite" p.307. Since equational sentences comprise a large part of the ISU relative clauses, the ISU position has shown a high tendency to have definite heads. Second, there is a difference between spoken and written data, since the data of these studies are obtained from spoken discourse, and Fox (1987, p. 858) suggests that relative clauses in spoken language tend to have more indefinite heads.

6.5.2 The Grammatical Position of the Head in the Main Clause

The position of the head of relative clauses in the main clause has some correlation with the frequency of relative clauses in the text, yet it does not determine that frequency. As is shown in section 5.4, noun phrases in different grammatical positions tend to be identified by ISU, TSU or DO relative clauses. For example, noun phrases that occupy the OBL and SU positions in the main clause tend to be identified by ISU relative clauses more than by relative clauses formed on other positions of the NPAH. Similarly, relative clauses whose heads occupy the GEN, SU(TOPIC), DO and SUCOM positions in the main clauses are mostly identified by DO relative clauses. The noun phrases in the category 'OTHERS' are the only positions that have the tendency to be identified by TSU relative clauses. 'OTHERS' have the lowest percentage among the grammatical positions of the head and this might lead to the result that TSU relative clauses do not have the highest frequency in the texts. The preference among noun phrases to be identified by ISU and DO more than TSU seems to support the AH. However, the results also suggest that the three upper positions, ISU, TSU and DO, are more accessible to relativisation than other relative clauses, because there are no instances found where a noun phrase has the highest tendency to be identified by OBL or GEN relative clauses.

Is the parallel function hypothesis proved by the data of this study?

Sheldon's (1974) claim that the relative clause is more accessible to comprehension when the head of the relative clause occupies the same grammatical position in the main clause and in the relative clause has not been supported by the data of this study. Heads of relative clauses in all grammatical positions tested tend to occupy the OBL position in the main clause. However, if only SU and DO positions are considered, as has been done by Sheldon, the results of this study show an agreement

with the results obtained by Sheldon. As is shown in section 5.4, DO-SU relative clauses, where SU and SU(TOPIC) are joined under SU, number 125; SU-SU relative clauses number 286; DO-DO relative clauses number 158; and SU-DO relative clauses number 283. Adding SU-SU and DO-DO results in a value of 444, which is slightly higher than SU-DO+DO-SU at 408. However, this cannot be extended to lower grammatical positions such as OBL and GEN. Counting the relative clauses where the functions of the heads in relative clauses parallel their functions in the main clauses, I have found that 688 relative clauses have heads with parallel functions; this value is lower than the value of other relative clauses where there is no parallelism (1809). Thus, the frequency of relative clauses in the texts suggest that the parallel hypothesis cannot explain the distribution of relative clauses other than SU and DO.

Regarding the relation between definiteness and the grammatical position of the head, the results show that heads in all grammatical positions tend to be definite except SU and OTHERS. OBL heads have the highest tendency to be definite (see section 5.4). This result is in agreement with previous results which state that there is a tendency to have OBL heads which are mainly definite because locative objects, which constitute a large part of OBLs, tend to be definite (Givón, 1979, pp. 53-54). Givón (1979) also states that the GEN position has a preference towards definiteness and this is also confirmed in this study, since the GEN position comes second after the OBL in the tendency towards definiteness. Noun phrases in the 'OTHERS' position tend to be indefinite and this is also in agreement with the results found by Fox and Thompson (1990), where existential heads, which are the biggest constituent (31%) of the 'OTHERS' category in this study, are found to be mainly indefinite. Thus, the results of this study confirm that the position of the head in the main clause helps in predicting the definiteness of the head.

The distribution of the grammatical positions of heads in the main clause positively correlates with the frequency of definite relative clauses. As explained before, heads in all grammatical positions tend to be definite except in two positions 'SU' and 'OTHERS'. The results show that heads occupying these two positions, SU and OTHERS, are infrequent (9% and 2%, respectively), which might have led to the small percentage of indefinite relative clause.

6.5.3 Animacy

The results show that the distribution of relative clauses that modify animate (i.e. human and nonhuman animate) referents is different from that of relative clauses that modify inanimate referents (see section 5.5). The NPAH is not reflected in any of the orders shown by the results of relative clauses that refer to animate and inanimate referents.

As far as the three upper positions, ISU, TSU and DO, of the hierarchy, the results suggest that the AH is manifested in the order of relative clauses that modify inanimate referents as is clear in their order of frequency DO, ISU, OBL>TSU>>GEN. On the other hand, the SPH is manifested in the order of relative clauses that identify animate referents, ISU, TSU>>DO>>GEN>>OBL. These changes can be attributed to the different tendencies among heads of relative clauses to be animate or inanimate. For example, DO heads have a high tendency to be inanimate whereas ISU heads have a high tendency to be animate. This is also found in previous studies such as Roland et.al. (2007) and Mak et.al. (2002). The latter suggest that inanimate is associated with DO, so having DO relative clauses with animate heads makes it harder to process than DO relative clauses with inanimate heads. This explains the infrequent use of DO relative

clauses with animate heads in the texts included in this study. The distribution of relative clauses that refer to human referents and those that refer to inanimate referents suggests that animacy has a significant influence on the distribution of relative clauses and on whether this distribution follows the NPAH or the AH.

The results of this study are found to be in agreement with the previous studies which state that animacy has an important role in the acquisition of relative clauses (e.g. Hawkins, 2007; Jeon & Kim, 2007; Kanno, 2007; Ozeki & Shirai, 2007). This study also conforms with previous corpus studies (e.g. Fox & Thompson, 1990; Mak et al., 2002; Reali & Christiansen, 2007) which state that animacy affects the frequency of SU and DO relative clauses. Although this study differs from the previous studies in that it considers the animacy of the head and does not consider the animacy of the noun phrases inside the relative clause, it provides similar results, in that animacy is found to be an influential factor that affects the distribution of relative clauses in a corpus.

The frequent use of inanimate heads in comparison with animate heads has also been found in previous language acquisition studies such as Diessel (2009) and Ozeki (2011), where children in both studies produced relative clauses with inanimate heads more than relative clauses with animate heads. The latter suggests that children tend to refer to animate referents by their proper names, so they do not need to use relative clauses to identify them, especially given that children tend to talk about close people in the surrounding environment. A similar explanation can be provided for the infrequent use of relative clauses with inanimate heads in the data of this study. The writers also tend to refer to persons by their proper names, so relative clauses are not necessary to identify them, and even if relative clauses are used in addition to their names, relative clauses will not be counted in the data of this study since they are not restrictive.

The results of the animacy data do not support the NPAH because, for the NPAH to be universal, the distribution of relative clauses in a text should follow the NPAH regardless of whether the head is animate or inanimate. The results do not support the AH either, as the AH suggests that ABS relative clauses are always preferred to SU relative clauses, which is found to be the case with relative clauses that modify inanimate referents, but not with those that identify animate referents. Furthermore, the distribution of ABS vs. SU relative clauses in relation to animacy might have led to the manifestation of the AH in the distribution of relative clauses, because the percentage of relative clauses whose heads refer to inanimate entities, where the AH is reflected, is significantly higher than the percentage of relative clauses whose heads refer to animate entities. Thus, animacy is found to have a significant influence on the distribution of relative clauses in Arabic discourse.

The relation between the animacy of the head and definiteness is not found to be significant (see section 5.5). Relative clauses that modify animate referents and those that modify inanimate referents both have a higher tendency to use definite heads, which is demonstrated in the three varieties of Arabic and in the two genres. Thus, the results of this study do not conform with the claim raised in previous studies (e.g. Givón, 1979; Lyons, 1999) that noun phrases that refer to human referents tend to be definite. Thus, the results suggest that the definiteness of the head of the relative clause is more related to its grammatical position in the main clause, as has been discussed in the previous section, rather than its animacy.

6.6 Summary

A discussion of the results shows that the order of the NPAH has not been reflected in the data of this study. The AH, as opposed to the SPH, is reflected in the overall results of this study. However, the overall results do not match the results of individual varieties and genres.

The distribution of relative clauses differs from one variety of Arabic to another and between fiction and nonfiction texts. The AH is manifested in CA rather than MSA and IA, and it is also reflected equally in both genres. These differences suggest that construction-external factors, especially variety, play an important role in the frequency of relative clauses, mainly relative clauses in the three upper positions: TSU, ISU and DO. DO relative clauses are preferred over ISU relative clauses in certain texts and vice versa.

The distribution of relative clauses has been explained in light of different factors such as definiteness, the position of the head in the main clause and animacy. Definiteness and the position of relative clauses in the main clause affect the distribution of relative clauses, yet these factors do not change the main result that the AH is supported more than the NPAH by the data of this study. On the other hand, the animacy of the head seems to determine whether relative clauses follow the AH or support the SPH. The AH is well manifested in the distribution of relative clauses that identify inanimate referents whereas the SPH is supported by the distribution of relative clauses that identify animate referents.

In conclusion, the discussion of the results of this study suggests that the distribution of relative clauses in written texts is determined by different factors. First, there is the accessibility of grammatical positions to relativisation, since relative clauses

in lower positions on the NPAH are infrequent regardless of the type of text they are used in. Second, the functions of relative clauses in discourse also play an important role in the distribution of relative clauses, especially if they are considered in relation to other factors such as genre and variety. Third, there are the characteristics of the head of the relative clauses such as definiteness, the grammatical position of the head in the main clause and animacy, which are called, in this study, construction-internal factors, which also have an important role in the frequencies of relative clauses. Some of these factors are related, such as definiteness and the grammatical position of the head, and animacy and the grammatical position of the head. Thus, none of these factors can alone determine the distribution of relative clauses in written texts.

Chapter Seven: The Pronoun Retention Strategy

7.1 Introduction

As discussed in Chapter One and Chapter Two, there are two relativisation strategies used in the Arabic language: the pronoun retention strategy, where the RP is used, and the gap strategy, where there are no RPs used. This chapter presents the results for the pronoun retention strategy in relative clauses found in Arabic texts. Different claims and theories on the use of the pronoun retention strategy have been discussed in the literature (section 3.3.2.2). The results in this chapter are presented according to these claims, starting with the NPAH in section 7.2. Section 7.3 presents the syntactic factors that might affect the use of RPs in relative clauses. Section 7.4 presents a discussion of how the use of the pronoun retention strategy might affect the pragmatic interpretation of the relative clause. The distribution of RPs across varieties of Arabic and genres is presented in section 7.5. Finally, a summary of the chapter is presented in section 7.6. The results of this chapter show that the NPAH provides the correct prediction for the use of the pronoun retention strategy (almost 95%). However, considering more than one factor, such as the distance between the head and the relativised noun phrase and head complexity, in addition to the grammatical position of the relativised noun phrase provides a better description of the use of the pronoun retention strategy vs. the gap strategy in Arabic relative clauses.

The data in this chapter is presented in raw percentages because the number of RPs is sometimes found to be too small in the SU position, for example, to be calculated using regression analysis. Counting RPs in relative clauses in different positions while considering different factors, I have obtained results with several zeros,

which is problematic in Poisson Regression models. Therefore, the data of the distribution of the RPs is presented in percentages.

7.2 The NPAH

According to the NPAH's predictions, the number of RPs increases going down the hierarchy since the grammatical positions that are lower on the NPAH are less accessible to relativisation than the upper grammatical positions. The use of RPs in the grammatical positions of the NPAH is shown in Figure 27.

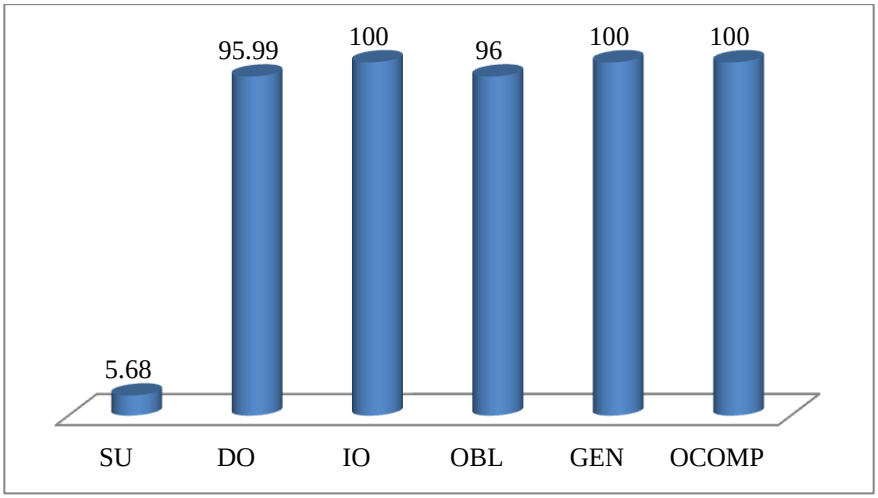


Figure 27: The percentage of relative clauses in each sentence type that includes RPs

As is shown in Figure 27, RPs are found in all the grammatical positions of the NPAH. There are three positions where the pronoun retention strategy shows variation; these positions are SU, DO and OBL. Relative clauses formed on IO, GEN and OCOMP positions always use the pronoun retention strategy. It is clear that the pronoun retention strategy is the preferred strategy in the DO position and other grammatical positions down the NPAH because of the big difference between the use

of RPs in the SU position and the use of RPs in other positions. The number of RPs found in the SU position is substantially less than the number of RPs in DO. The OBL position has a slightly higher percentage of RPs than the percentage of RPs in DO position, yet less than the percentage of RPs in IO position.

Examples (95-103) show the use of RPs in Arabic relative clauses in different grammatical positions of the NPAH starting with the SU position, where examples (95) and (96) show the SU relative clauses without and with the RP, respectively. Then, examples (97) and (98) show the DO relative clauses without and with the RP, respectively. Examples (100) and (101) show the OBL relative clauses without and with the RP, respectively. Examples (99), (102) and (103) show IO, GEN and OCOM relative clauses, respectively.

SU

95. ʾinna al-ağsāma allatī taqbal Ø al-iḍāʾah ….
indeed the-bodies REL(3.SG.FEM) she.acceptØ the-light…
Indeed, the bodies which receive the light…
[CA, 3: 10]

96. al-faʾl allaḏī *huwa* yurīdu al-ẖair
the-augury REL(3.SG.MAS) *he* he.want the-good
The augury which bodes the good things.
[CA, 2: 399]

DO

97. man al-raǧul allaḏī raʾai-tu-Ø maʿa-k
who the-man REL(3.SG.MAS) saw-I-Ø with-you
Who is the man whom I saw with you?
[CA, 4: 84]

98. ʾaḥbab-tu ʾan lā ʾansarif ḥattā yatabayan l-i
loved-I to not go.I until clear for-me
sabab al-ḍawʾ allaḏī raʾi-tu-*hu*
reason the-light REL(3.MAS.SG) saw-I-*it*
I did not want to go until I knew the reason behind the light which I saw.
[CA, 1: 255]

IO

99. kān ʾawal wāli faraḍ la-*hu* raʿiyat-hu
was first magistrate assigned for-*him* people-his
nafaqāta-hu
expenses-his
He was the first magistrate to whom people assigned his expenses.
[CA, 5:306]

OBL

100. fī al-waqt allaḏī qāmat Ø-Ø bi-l-taʿtīm ʿala
in the-time REL(3.SG.MAS) did Ø-Ø in-the-hiding on
ruwat al- qismi al-ṯānī
narrators the-part the-second
The time in which they have hidden the narrators of the second part.
[MSA, 5: 168].

GEN

101. taḥta-hu baġlat la-hu qad quṭiʿa
under-him mule-FEM for-him indeed cut-PASS
ḏanabu-*hā*
tail-*its*
He was on his female mule whose tail was cut.
[CA, 4: 1107]

OCOMP

102. qad nazala bi-nā qaum lam narā
already came down in-us people not we.see
qawm qaṭ ʾaḥsan min-*hum*
people never better than-*them*
We have never seen people better than the people who came to our house.
[CA. 4: 103]

Using only the pronoun retention strategy in the GEN and OCOMP position supports the NPAH. On the other hand, using only the pronoun retention strategy in the IO position might suggest that this strategy does not apply to a continuous segment of the NPAH since the pronoun retention strategy is not the only strategy used in the DO and OBL positions. However, the suggestion that RPs in Arabic does not apply to a continuous segment of the NPAH cannot be confirmed in this study because of the small number of IO relative clauses in comparison with relative clauses in other positions. In fact, in Arabic, the use of RPs is not obligatory in IO relative clauses unless the pronouns are attached to prepositions (Al-dahdah, 2007; Al-muradi, 2001). There are only two IO relative clauses found in the data of this study, and in these two cases, RPs are attached to the prepositions, as is shown in example (99). The small number of IO relative clauses found does not help to show different strategies,

especially if the pronoun retention strategy is the preferred strategy in DO and other positions down the NPAH.

7.2.1 NPAH's Claims on the Arabic Language and Violations to These Claims

Keenan and Comrie claim that, in the Arabic language, the pronoun retention strategy is used in all of the grammatical positions of the NPAH except in the SU position (1977, pp. 76, 93; 1979, p. 333). However, in this study, some violations to this claim are found.

The violations discovered relate to the use of the pronoun retention strategy in grammatical positions where RPs are predicted to be absent (i.e. the subject position), and the absence of RPs in positions where these pronouns are predicted to be retained (DO and other positions down the NPAH). The results show 128 violations, including 1 TSU relative clause, 78 ISU relative clauses, 30 DO relative clauses and 19 OBL relative clauses, as is shown in Table 24 below.

Table 25: Violations to the Predicted Distribution of RPs in Relative Clauses in Arabic Texts

The position of violations	The number of relative clauses in each position	The number of violations
TSU	601	1 (0.16%)
ISU	789	78(9.88)
SU	1390	79(5.68%)
DO	749	30 (4%)
OBL	475	19 (4%)
Total	2614	128 (4.89%)

In fact, the highest percentage of violations is found in the SU position, which is surprising. Having a look at the SU relative clauses where RPs are used, I found that 64 relative clauses, 81% of SU relative clauses, have prepositional verbs in the passive voice, as is shown in example (103).

103.	kān	fī	masǧid	yuqāl	la-*hu*	masǧid
	was	in	mosque	say(PASS)	to-*it*	mosque
	ṣaliḥ					
	Salih					

He was in a mosque which is called Salih's mosque.
[CA, 4: 80]

As is shown in example (103), the RP is attached to the preposition *la* 'to'. Although the pronoun has the form of an object of preposition, which makes the relative clause appear as an OBL relative clause, it is considered as an ISU relative clause for two reasons. First, the prepositional verb is considered a transitive verb because the prepositional phrase *la-hu* 'to-it' is semantically attached to the verb *yuqāl* 'said' and cannot be separated (see section 4.4.1.1); second, the verb is in the passive form. In this case, the omission of the pronoun is governed by the rule of preposition stranding, which is not allowed in the Arabic language (Moutaouakil, 1989, p. 65). Therefore, whenever the preposition is used, the object of preposition, which is the RP, should be used too. In the data of this study, there are no prepositional verbs used without the use of RPs.

Other SU relative clauses that use RPs are distributed as follows: 1 TSU (see example (96), above) and 8 ISU relative clauses including 6 equational sentences and 2 prepositional phrases, as shown in examples (104) and (105), respectively:

104. ṣaraḫ ʿala al-ḫādim allaḏī huwa nāʾim
shout.he at the-servant REL(3.SG.MAS) he asleep
He shouts at the servant who is asleep.
[CA, 1: 103]

105. wağad niṣfa-hā allaḏī huwa fī al-ğānib
found.he half-her REL(3.SG.MAS) he in the-side
al-wāḥid
the-one
He found her half which is on one side.
[CA, 3: 34]

The use of RPs in SU relative clauses is supported by previous studies (e.g. Jassim, 2011; Suaieh, 1980) which claim that RPs can be used in SU relative clauses for emphatic reasons.

RPs are not used in 30 DO relative clauses (4% of the total number of DO relative clauses). Suaieh (1980) claims that relative clauses with RPs are well formed, yet the grammatical rules of Arabic do not the prohibit the absence of RPs in DO relative clauses (Al-dahdah, 2007; Al-muradi, 2001)

RPs are not used in 19 OBL relative clauses (4% of the total number of OBL relative clauses). In all these sentences, the whole prepositional phrase is absent and not only the RPs, consider example (100) above. In example (100), the prepositional phrase *fi-hi* ‘in-it’, which is supposed to appear instead of the symbols Ø-Ø, is absent. Because preposition stranding is not allowed in the Arabic language, it is possible for the RP not to be used on the condition that the preposition is not used either. Since the preposition is not semantically attached to the verb in the case of the OBL relative clause, and the absence of the pronoun will not cause a difference in meaning, it is possible not to use the prepositional phrase including the RP.

The violations found are only 4.89% of the total number of relative clauses, and the percentages of violations in positions shown in Table 25 do not exceed 6% in each position, if we count the TSU and ISU as one position, the SU position, as is found in the NPAH. It follows that the NPAH provides the correct predictions for the use of RPs for around 95% of the relative clauses in the corpus. Thus, the use of the pronoun retention strategy in Arabic texts provides evidence for the claim that the use of RPs simplifies the processing of difficult relative clauses, which are found in the least accessible positions on the NPAH (Hawkins, 1999a; Keenan & Comrie, 1977, 1979).

7.3 Factors that Affect the Use of RPs

In this section, I evaluate some of the theories that have tried to explain the use of RPs in relative clauses. These theories associate the use of the pronoun retention strategy with factors such as the distance between the head and the relativised noun phrase (see section 7.3.1), definiteness (see section 7.3.2) and animacy (see section 7.3.3).

7.3.1 The Distance between the Head and the Relativised Noun Phrase

The use of the RP has been explained in relation to the processing of the relative clause. One of the processing difficulties arises from the distance between the head and the relativised noun phrases. A number of studies (e.g. Ariel, 1990; Collier, 1991; Prince, 1990) claim that the greater the distance between the head and the position of the relativised noun phrase is, the more difficult the relative clause is to process, and the greater the tendency is to use an RP in that relative clause.

The effect of the distance between the head and the relativised noun phrase is tested in this study. The distance between the head and the relativised noun phrase is

given a score according to the number of words between these two elements in the sentence while ignoring the relative marker[29]. Using the same scheme used by Ariel (1999), a score of 0 is assigned if no words or one word occur between the head and the relativised noun phrase; 1 if two words occur, and 2 if three words or more occur, as is shown in sentences (106), (107), (108), respectively. In these examples, to show the number of the words between the head and the relativised noun phrase, these words are underlined.

106.	ʾinna	al-aǧsām	allatī	taqbalØ	al-ʾiḍāʾah…
	indeed	the-bodies	REL(3.SG.FEM)	she.acceptØ	the-light…

Indeed, the bodies which receive the light… [CA, 3: 10]

107.	aqraʾu	ʿalai-hā	qaṣāʾid-ī	allatī	lam
	read(I)	on-her	poems-my	REL(3.SG.FEM)	not
	taqraʾ-*hā*	fī	ḥayāti-hā		
	read.she-*it*	in	life-her		

I read to her my poems that she did not read in her life. [MSA, 5:144]

108.	ḫarağ	ʾahlu	al-Kufah	allaḏīn	kān
	march out	people	the-Kufah	REL[3.PL.MAS]	was
	Al-Muḫtār	qātala-*hum*			
	Al-Muḫtar	fought-*them*			

The people of Kufa whom Al-Muḫtar fought marched out. [CA, 4:1107]

The results show that the percentage of relative clauses using RPs is very low (29.5%) when the score is zero (i.e. there is only one word or no words between the

[29] In Arabic, definite relative clauses use relative markers while indefinite relative clauses do not. Counting the relative marker as a word will make definite relative clauses have more words than indefinite relative clauses. In order to eliminate the factor of definiteness the relative marker is not counted as a word.

head and the relativised noun phrase). On the other hand, the percentage of relative clauses using RPs when the score is one (94.01%) or two (94.02%) is very high with a slight difference between the two, as is shown in Figure 28.

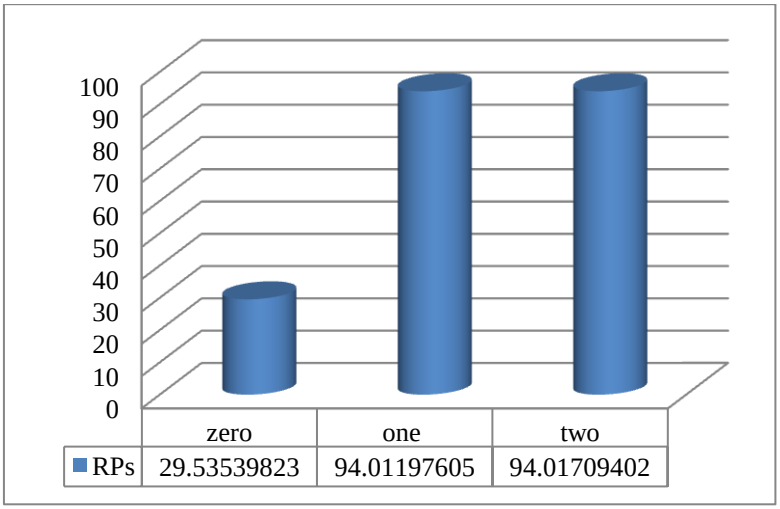


Figure 28: The percentage of relative clauses using the RPs in relation to the distance between the head and the relativised noun phrase

When only one word or no words occur between the head of the relative clause and the relativised noun phrase, there is a tendency to use the gap strategy. However, the tendency towards using the pronoun retention strategy is similar when two, three or more words are used between the head of the relative clause and the relativised noun phrase. Thus, the results of this study support the claim that the distance between the head of the relative clause and relativised noun phrase affects the use of the pronoun retention strategy.

7.3.2 Definiteness

A number of studies (e.g. Prince, 1990; Tarallo, 1986) show that the RP is preferred more when the head is indefinite than when the head is definite. Previous

studies on the Arabic language such as Galal (2004), Suaieh (1980), and Shlonsky (1992) state that RPs are required in indefinite relative clauses (i.e. relative clauses that modify indefinite heads) (for more details on this see sections 3.3.2.1 and 3.3.2.2).

The results of this study show that there is a slightly higher preference towards using the pronoun retention strategy in relative clauses whose heads are indefinite. Indefinite relative clauses with RPs have a higher percentage (14.98%) than indefinite relative clauses using gaps (12.36%). On the other hand, definite relative clauses have similar tendencies towards using RPs and gaps since the percentage of definite relative clauses using the pronoun retention strategy (36.12%) is so close to the percentage of definite relative clauses using the gap strategy (36.52%), as is shown in Figure 29.

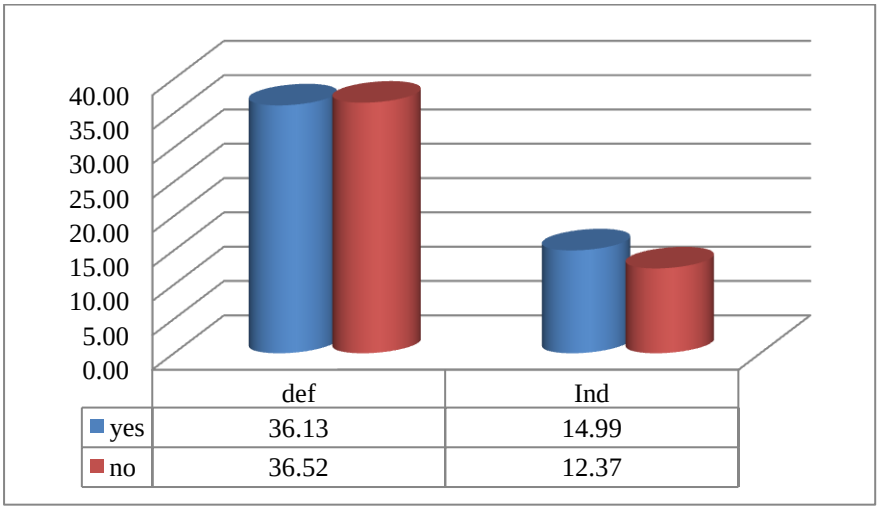


Figure 29: A comparison of the percentages of (def.) definite and (ind.) indefinite relative clauses with and without RPs, where 'yes' indicates the use of RPs, and 'no' indicates the use of gaps

The tendency to use RPs in relative clauses whose heads are indefinite is found true especially in DO relative clauses. In fact, the data of this study show only one DO

relative clause where the head is indefinite and the RP is absent, as is shown in example (109).

109.	fa	ṣallaṭa	ʿalai-hum	ḥarran	la	yastaṭīʿ-ūn	an	
	so	brought(he)	on-them	hot	not	can-PL	to	
	yataqār-ū-Ø							
	bear-PL-Ø							
	So HE (Allah) brought to them hot weather which they could not put up with. [CA, 4: 111]							

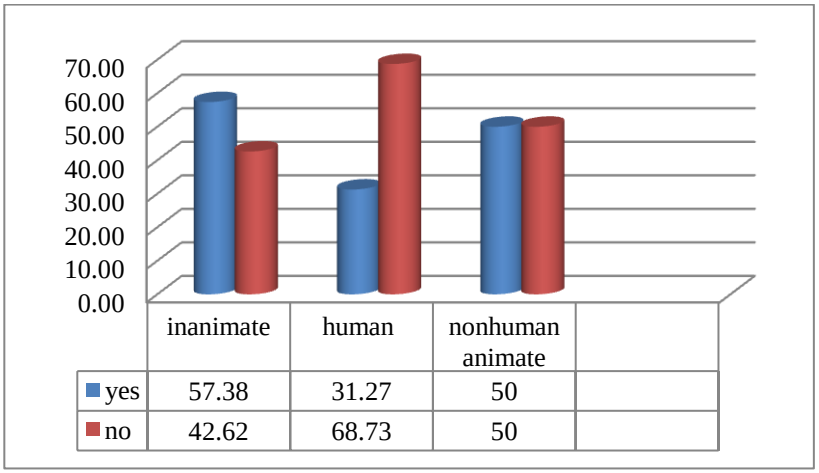
Suaieh (1980) argues that, in Arabic, the relative markers help the reader to infer the RP if it is absent because they are inflected for number and gender. On the other hand, the absence of the relative marker in indefinite relative clauses, since, in Arabic, relative markers are used in definite relative clauses only, makes it difficult for the reader to identify the RP. Accordingly, the absence of RPs in indefinite relative clauses might cause difficulty in text comprehension.

7.3.3 Animacy of the Head

Erteschik-Shir (1992) noted that the use of RPs is preferred with heads that refer to human or animate referents (pp. 104-105). This has been supported by studies on different languages (e.g. Bošković, 2009: on Serbo-Croatian; Jaggar, 2001: on Hausa; Salzmann, 2006: on Zurich German) It follows that the animacy of the head might influence the use of RPs.

The results of this study show that relative clauses whose heads refer to inanimate referents have a higher tendency towards using the pronoun retention strategy. Relative clauses whose heads refer to nonhuman animate referents show the same tendency towards using the pronoun retention strategy or the gap strategy as

relative clauses with heads refer to inanimate. On the other hand, relative clauses that identify animate referents have a preference towards using the gap strategy, as is shown in Figure 30.



	inanimate	human	nonhuman animate	
yes	57.38	31.27	50	
no	42.62	68.73	50	

Figure 30: A comparison of the percentages of inanimate, human and nonhuman animate relative clauses with and without RPs, where 'yes' indicates the use of RPs and 'no' indicates the use of gaps

As is shown in Figure 47, the percentage of relative clauses whose heads refer to human referents when using the gap strategy is 68.73%, which is higher than the percentage of relative clauses whose heads refer to human referents when using the pronoun retention strategy (31.27%), which is counter to previous claims. One possible explanation for having fewer RPs in relative clauses whose heads refer to human referents is that SU relative clauses, which mainly identify human referents, do not usually use RPs (see section 5.5).

Including relative clauses on all grammatical positions on the NPAH except the SU position changes the results, as is shown in Figure 31.

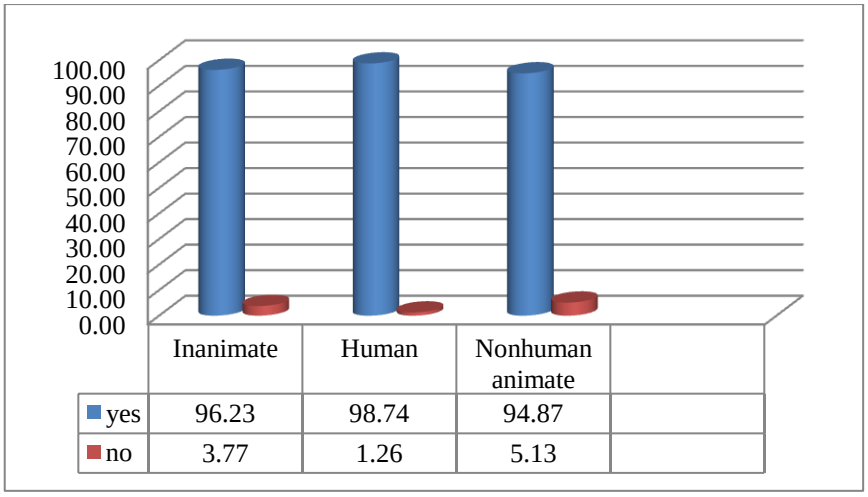


Figure 31: A comparison of the percentages of inanimate, human and nonhuman animate relative clauses, other than SU relative clauses, with and without RPs, where 'yes' indicates the use of RPs and 'no' indicates the use of gaps

As is shown in Figure 31, the three types of relative clauses, relative clauses refer to inanimate, human and nonhuman animate referents, have high tendencies towards using RPs; it is not only relative clauses whose heads refer to human referents that use more RPs. Thus, the claim that RPs are preferred more if the head of the relative clause refers to an animate referent is not proved in this study. One possible explanation for these results is that the pronoun system in Arabic shows a gender and case distinction rather than an animacy distinction (see section 3.2.1.3). Similar to animate entities, inanimate referents are distinguished into feminine and masculine. Since animate and inanimate entities are treated similarly, animacy is not found to be an important factor in the use of RPs in Arabic.

7.3.4 The Accessibility Theory

Ariel (1990, 1999) suggests that the use of the RP is affected by the degree of accessibility associated with the head. The degree of accessibility is determined by several factors. These factors are the distance between the head and the relativised noun

phrase, the restrictiveness of the relative clause, the complexity of the head and the obligatoriness of the relativised noun phrase. In this study, the accessibility claim is tested. The accessibility associated with the head is determined by counting the factors suggested by Ariel (1999) except that one factor, the restrictiveness of the relative clauses, is excluded from examination in this study, since all relative clauses included in this study are restrictive. Other factors are counted using the same method as Ariel (1999), which is as follows:

1. The distance between the head and the relativised noun phrase: 0 if one or zero word; 1 if two words, 2 if 3 words and above.
2. The head complexity: 0 if the heads consists of one word; 1 if two words, 2 if 3 words and above.
3. Obligatory argument: 0 if the relativised noun phrase is +obligatory, and 1 if it is –obligatory.

The value of these factors is counted in each sentence, and as a result, the heads associated with high accessibility receive a low score. For example, the accessibility associated with the head in example (110) is counted as follows:

a. (0) is given for the distance factor because there is only one word *tarġab* 'want(you)';
b. (0) is given for the head complexity because the head '*al-makān*' 'the-place' consists of one word in Arabic;
c. (0) is given for the third factor 'obligatoriness of the relativised noun phrase' because DO is an obligatory argument.

Therefore, the score will be zero, which makes the head highly accessible.

110. taṣrifa-nī matā arad-ta wa ʾilā al-makān
you(dismiss-me) when want-you and to the-place
allaḏī tarġab
REL(3.SG.MAS) want.you
You dismiss me whenever you want and to the place that you want.
[MSA, 3:152]

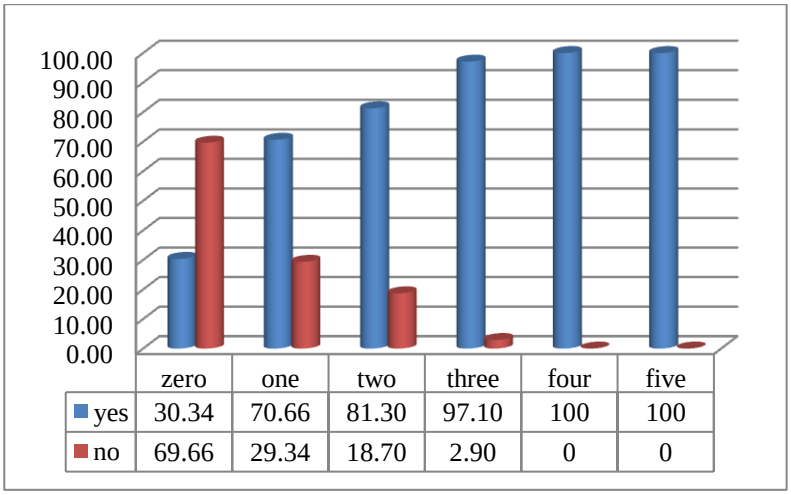
On the other hand, the accessibility associated with the head is low in (111):

111. aḍfā fuqahāʾ al-qaum al-mašrūʿīya wa al-
added jurists the-people the-legitmacy and the-
qadāsah ʿala al-qurūn al-ṯalāṯah al-ʾūlā
holiness on the-centuries the-three the-first
allatī našaʾa wa qunnina fī-*hā* al-fikr
REL(3.SG.FEM) arose and legalise.PASS in-*it* the-thought
al-Islamī
the-Islamic
Jurists added legitimacy and holiness in the first three centuries in which Islamic thought arose and was legalized.
[MSA, 5: 191]

In (111), the accessibility score is five; it can be calculated in the same way, the head consists of three words *al-qurūn al-ṯalāṯah al-ʾūlā* 'the first three centuries'. There are 4 words between the relativised noun phrase and the RP- *ha* 'it'- *naša'a wa qunnina fī* 'arose and legalized in'. The relativised noun phrase is an object of the preposition, which is not an obligatory argument. In this sentence, the score is high so an RP is used.

The results of this study show that the higher the score the relative clause receives (i.e. the lower the degree of accessibility associated with the head is), the more frequently RPs are used. The percentage of RPs increases with the score of two and

three until it reaches 100% with the scores of four and five (see example (111)). The score of zero is the only score related to a preference for the gap strategy (see example (110)), as is shown in Figure 32.



	zero	one	two	three	four	five
yes	30.34	70.66	81.30	97.10	100	100
no	69.66	29.34	18.70	2.90	0	0

Figure 32: A comparison of the percentages of relative clauses whose scores are from zero - five with and without RPs, where 'yes' indicates the use of RPs and 'no' indicates the use of gaps

Thus, the results of this study strongly support Ariel's (1999) claim that considering more than one factor better predicts the use of the pronoun retention strategy.

7.3.5 The RP as the Last Option

Shlonsky (1992) claims that RPs are found to be the last option for sentences which are otherwise ungrammatical. He claims that the use of the RP is only licensed when the gap yields an ungrammatical sentence by violating the grammar constraints of the language (see section 3.3.2.2). For example, he argues that RPs are obligatory in OBL relative clauses in Hebrew and Palestanian Arabic since both languages have a

constraint against preposition stranding (p. 446). However, this study yields 19 OBL relative clauses where the gaps are used instead of RPs.

He also explains that there is no true optionality between the use of RPs and the gap strategy in relative clauses, and that relative clauses which appear to be using the RP one time and the gap in another are morphologically similar, but not identical. He states:

> "there is no free variation between gaps and RPs; rather, there is an illusion of free variation that is propped up by the morphological nondistinctness of two crucially different types of relative clauses" (Shlonsky, 1992, p. 448)

However, the results of this study show that RPs are sometimes used in sentences which also use the gap strategy, as in examples (112) and (113) .

112.	bi-l-ʾisnādi	allaḏī	ḏakar-tu-hu	qabl
	in-the-reference	REL(3.MAS.SG)	mentioned-I-it	before
	In the reference which I mentioned before.			
	[CA, 4: 106]			

113.	bi-l-ʾisnādi	allaḏī	ḏakar-tu-Ø	qabl
	in-the-reference	REL(3.MAS.SG)	mentioned-I-Ø	before
	In the reference which I mentioned before.			
	[CA, 4: 1261]			

Examples (112) and (113) are similar except in the use of the RP as the same sentence uses an RP in (112) and a gap in (113). If the claim that the RP is only licensed when using the gap strategy violates the grammatical rules of the language is correct, then how could the RP be used in (112) while the gap is similarly grammatical as is shown in (113)? Shlonsky believes that there is no choice between having an RP

or a gap. However, "the choice seems to lie in the phonetic realization of the resumptive pronoun itself" (1992, p. 458). That is, example (113) has an RP similar to example (112), yet this RP is phonetically not realised. This claim cannot be proven since there is no observable sign that helps to ensure the presence of the RP in example (113).

Thus, the results of this study provide partial support for Shlonsky's (1992) claim in the case of the prepositional verbs. However, the results yield some examples, such as examples (112) and (113), where the RP and the gap appear to be in free variation. The theme of the last option, or what Shlonsky calls the 'last resort' cannot provide an adequate explanation for the use of the pronoun retention strategy in Arabic relative clauses.

7.3.6 Studies on Arabic

Arabic grammar books state that RPs can be used in all grammatical positions. Arab grammarians set up some rules that describe the conditions under which the RP can be optional (Al-muradi, 2001). In general, there are two conditions that had been put forward by traditional Arab Grammarians for the gap to be used instead of the RP: (1) if the gap does not cause any ambiguity; (2) if the RP can be recovered from the sentence. In this section, I will present the details of these two conditions.

Traditional Arab grammarians state that an RP must be used in the SU position if it is followed by either a 'prepositional phrase' or what they call *ṣilah kāmilah* 'a full relative clause', which means that it consists of the elements of a full sentence V.S.O (Al-dahdah, 2007; Al-muradi, 2001), as in examples (114) and (115) respectively.

114.	wağada	nişfa-hā	allaḏī	huwa	fī	al-ğānib
	found.he	half-her	REL(3.SG.MAS)	he	in	the-side
	al-wāḥid					
	the-one					
	He found her half which is on one side.					
	[CA, 3: 34]					

115.	al-faʾl	allaḏī	*huwa*	yurīd	al-ẖair
	the-augury	REL(3.SG.MAS)	*it*	want(it)	the-good
	The augury which bodes the good things.				
	[CA, 2: 399]				

In example (115), the RP is followed by what the traditional Arab grammarians call *ṣilah kāmilah* 'full relative clause'- *yurīd al-ẖair* "it wants the good things". The RP is the overt subject[30] of the relative clause. Al-muradi (2001) explains that, when the RP which is followed by a full relative clause is absent in the sentence, there is no *dalīl* 'sign' that leads the reader to recognize the existence of the RP because the relative clause is syntactically complete without the use of that RP. In some sentences, RPs are followed by incomplete relative clauses, e.g. followed by a one-word predicate, as in example (116).

116.	ṣaraẖa	ʿala	al-ẖādim	allaḏī	huwa	nāʾim
	shout	at	the-servant	REL(3.SG.MAS)	he	asleep
	He shouts at the servant who is asleep.					
	[CA, 1: 103]					

In example (116), the relative clause is an equational sentence; it consists of the RP, which occupies the subject position, followed by the predicate, *nā'em* 'asleep'. In

[30] Arabic is one of the null subject languages where the subject can be understood because the verb is highly inflected (Ouhalla, 1999; Ryding, 2005, p. 299) (see section 3.2.1.3).

such relative clauses, the RP occupies an obligatory argument slot because it is the *mubtadaʾ* 'subject', so when it is absent the sentence is incomplete. Therefore, there is a 'sign' *dalīl* that leads the reader to recognise that there was an RP which has been deleted (Al-muradi, 2001, pp. 451-452).

According to traditional Arab grammarians, an RP is optional in relative clauses such as (116). On the other hand, an RP must be used if it is followed by a full sentence such as example (115), or if it is followed by a prepositional phrase such as example (114).

Applying this rule to the data of this study, I have found that only seven of the SU relative clauses (9.5%) that use the RP are full relative clauses, and two of them (2.7%) are prepositional phrases, while the other 68 SU relative clauses (86.07%) use prepositional verbs inside the relative clauses which, the Arab grammarians did not account for. Moreover, the grammatical rules that have been set by traditional Arab grammarians do not permit the choice between an RP and zero in sentences where the relative clause consists of either 'a full relative clause', 'prepositional phrase' or, 'adverbial clause' (Al-dahdah, 2007, p. 64; Al-muradi, 2001, p. 452). However, the data of this study show some counter examples to this rule, as is shown in examples (117) and (118):

117.	ʾinna	al-ağsāma	allatī	taqbal	al-ʾiḍāʾeh ….
	indeed	the-bodies	REL(3.SG.FEM)	she.accept	the-light…
	Indeed, the bodies which receive the light……				
	[CA, 3: 10]				

118.	qaṭṭaʿ-nna	aidīya-hunna	bi-l-sakākīn	allatī	fī
	cut-PL.FEM	hands-their	by-the-knives	REL(3.SG.FEM)	in
	aidīya-hunn				
	hands-their.FEM				
	They cut their hands with the knives which were in their hands.				
	[CA, 4: 116]				

In sentence (117), the relative clause is *ṣilah kamilah* 'a full relative clause'- *taqbal al-ʾiḍāʾeh* 'they accept the light', and in (118) the relative clauses is a prepositional phrase- *fi ayīya-hunna* 'in their hands', yet the two relative clauses are formed without the use of RPs. Following the grammatical rules for the use of RPs in SU relative clauses, these two examples and similar sentences are considered ungrammatical. Thus, the rules set out by traditional Arab grammarians cannot adequately describe the use of RPs in SU relative clauses.

RPs are considered optional in DO relative clauses provided that the absence of RPs does not cause any ambiguity in the text (i.e. the RP is recoverable) (Al-muradi, 2001; Galal, 2004). Consider the invented examples (119) and (120).

119.	raʾai-tu	al-šurṭi-**a**	allaḏī	ḍaraba
	saw-I	the- police officer-**ACC**	RE(3.SG.MAS)	hit
	al-raǧul-**u.**			
	the-man-**NOM.**			
	I saw the police officer whom the man hit.			

120.	raʾai-tu	al-šurṭi-**a**	allaḏī	ḍaraba
	saw-I	the-police officer-**ACC**	REL(3.SG.MAS)	hit
	al-raǧul-**a**.			
	the-man-**ACC**.			
	I saw the police officer who hit the man.			

The ambiguity occurs when the case marker is not used on both the head *al-šurṭi* 'the-police officer' and the subject or the object of the relative clauses *al-raǧul* 'the-man', as is shown in example (121), where it is difficult to judge whether the head is the subject (the agent) or the object (the patient) of the relative clause .

121.	raʾai-tu	al-šurṭi	allaḏī	ḍaraba	al-raǧul
	saw-I	the- police officer	RE(3.SG.MAS)	hit	the-man
	I saw the police officer whom the man hit.				

Case markers are dropped in most texts except in religious books (Galal, 2004). Therefore, the absence of the RP mostly causes an ambiguity in examples similar to (121). However, if the subject of the relative clause is a pronoun, then no ambiguity arises, since most pronouns are marked for grammatical case, as there are subject and object pronouns (Al-dahdah, 2007; refer to section 2.4.1.3). In example (121), if the subject of the verb in the relative clause was a pronoun like *–tu* 'I', it would be easier to recognise that the relative clause is a DO relative clause as the head is the DO of the relative clause. This would predict that the subject position in DO relative clauses whose RPs are absent is mostly occupied by pronouns[31] rather than full noun phrases.

The results of the present study show that pronouns occupy the subject position in 23 DO relative clauses with no RPs, which make up 82% of all DO relative clauses where RPs are not used, as is shown in Figure 33.

[31] It is not necessary that the subjects are overt pronouns since Arabic is a null subject language, the subject pronoun can be understood from the form of the verb.

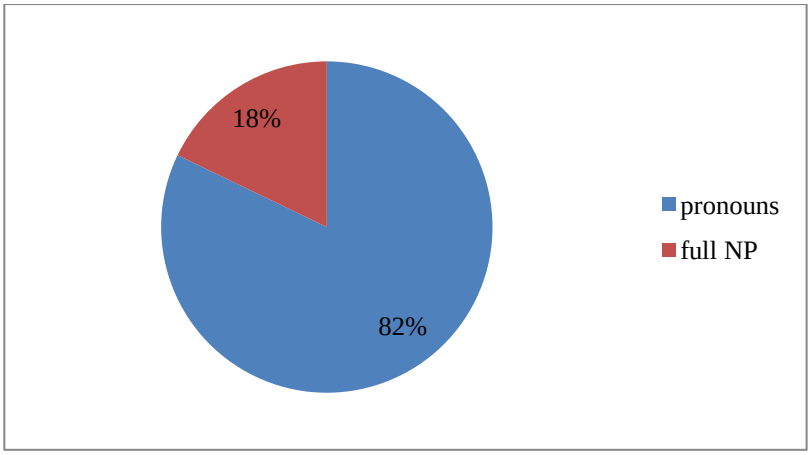


Figure 33: Percentages of pronouns vs. full noun phrases that function as SU in DO relative clauses which use the gap strategy

Different kinds of pronouns- first, second and third person pronouns are used as the subjects of DO relative clauses, as exemplified in sentences (122), (123) and (124), respectively.

122.	man	al-rağulu	allaḏī	raʾai-tu	maʿa-k
	who	the-man	REL(3.SG.MAS)	saw-I	with-you

Who is the man whom I saw with you? [CA, 4: 84]

123.	taṣrifa-nī	matā	arad-ta	wa	ʾilā	al-makān
	you(dismiss-me)	when	want-you	and	to	the-place
	allaḏī	tarġab				
	REL(3.SG.MAS)	want.you				

You dismiss me whenever you want and to the place that you want. [MSA, 3:152]

124.	fa-halak-ū	fī	al-maǧāʿah	allatī	ṣādaf-
	so-died-PL	in	the-famine	REL(3.SG.FEM)	they.found-
	ū				
	PL				

So they died in the famine which they suffered from. [CA, 3:1000]

The results show that when pronouns occupy the subject position inside relative clauses, there is a slightly higher tendency to use the gap strategy, as is shown in Figure 34.

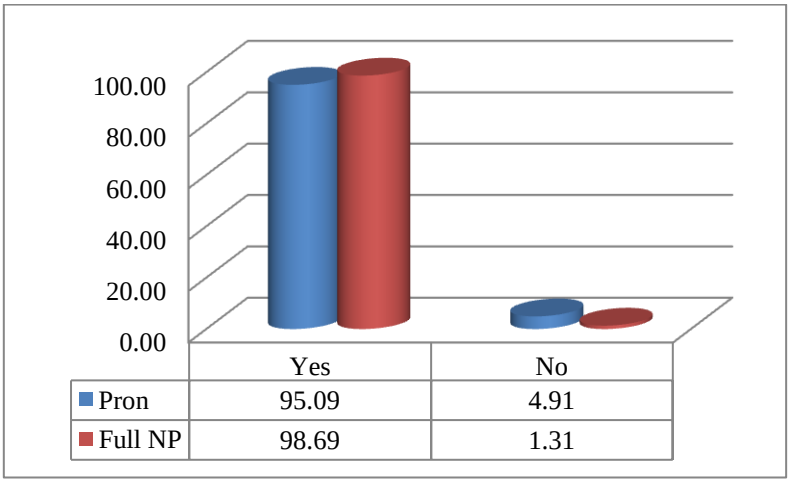


Figure 34: A comparison between the percentages of relative clauses with and without RPs when the subject position inside the relative clause is occupied by pronouns and full noun phrases, 'yes' indicates the use of RPs and 'no' indicates the use of gaps

As is shown in Figure 34, the percentages of relative clauses that use the pronoun retention strategy is high in both cases when the subject position inside the relative clause is occupied by a pronoun (95.08%) or a full noun phrase (98.68%). However, most of the relative clauses that use the gap strategy use pronouns (4.91%) rather than full noun phrases (1.31%) as their subjects.

Al-dahdah (2007) and Bashir (1982) claim that, in OBL relative clauses, the preposition together with the RP are deleted when the RP is preceded (governed) by the same preposition which precedes the head. It follows that it is necessary for the head of the relative clause to be oblique. However, results have shown that only nine relative clauses out of nineteen (47.36%) have OBL heads. This indicates that using the same

preposition before the head and before the RP is not a necessary condition for the absence of the prepositional phrase.

Badawi (2004) claims that the prepositional phrase in the relative clause can be absent if the head is a time expression. It follows that the head should be either an adverb of time or an object of preposition of time. The results show that sixteen out of nineteen relative clauses (84.21%) whose prepositional phrases are absent have time expressions as their heads such as *waqt* 'time', *yawm* ' a day', *lailah* 'a night' and *lahẓa* 'a second'. An explanation for this is provided by Givón (1993), who, based on the grammar of English, states:

> "REL-clauses that involve certain adverbial roles – time, manner, reason, purpose, or even place — have some unique characteristics. They often employ the gap strategy of relativisation, with no concomitant problems of case-role recoverability. This is so because either the head noun or the REL-pronoun in such constructions is semantically specific revealing through its meaning the likely case-role of the missing co-referent noun" (pp. 135-136).

Givón (1993) provides different examples of different relative clauses and adverbials among which are the following:

- The time he was supposed to meet me...
- When she finally showed up... (p. 136).

Although, in Arabic, there are no specific relative markers for time or place as in English 'when' and 'where', the head in the Arabic OBL relative clause, similar to the case of English, is so specific that it is enough to recover the missing pronoun as in example (100), which is repeated here as (125).

125. ...fī **al-waqti** allaḏī qāmat Ø-Ø bi-l-taʿtīmi
...in **the-time** REL(3.SG.MAS) did Ø-Ø in-the-hiding
ʿala ruwat al-qism al-ṯānī
on narrators the-part the-second
…the time in which they have hidden the narrators of the second part.
[MSA, 5: 168].

To summarise, studies on Arabic present different rules for the use and absence of RPs in different grammatical positions. These rules have one shared point which is that the RP, if not used, should be recoverable, and the absence of RPs should not cause any ambiguity in the sentence.

7.3.7 Summary

Results of this study show that the use of RPs is mostly predicted by the NPAH, yet the prohibition of preposition stranding in Arabic interferes with the NPAH's predictions. It has also been found that certain factors do affect the distribution of RPs in relative clauses. Among these factors are definiteness and the distance between the head and the relativised noun phrase. RPs are preferred when the relative clause identifies an indefinite head, and the longer the distance between the head and the relativised noun phrase is, the more likely the RP is to be used. The results of this study also show that calculating the accessibility associated with the head by counting different factors at the same time provides the correct prediction for the use of the pronoun retention strategy; the lower the degree of the accessibility associated with the head the more likely RPs are used.

7.4 The Interpretations of Relative Clauses with and without RPs

The results of this study show some examples where there are no syntactic or semantic differences, apart from the use of the RP, between the DO relative clauses that retain pronouns and those which do not, consider the examples (112) and (113), which are repeated here as (126) and (127):

126.	bi-l-ʾisnādi	allaḏī	ḏakar-tu-hu	qabl
	in-the-reference	REL(3.MAS.SG)	mentioned-I-it	before
	In the reference which I mentioned before.			
	[CA, 4: 106]			

127.	bi-l-ʾisnādi	allaḏī	ḏakar-tu-Ø	qabl
	in-the-reference	REL(3.MAS.SG)	mentioned-I-Ø	before
	In the reference which I mentioned before.			
	[CA, 4: 1261]			

A number of studies that have discussed the use of RPs in relative clauses consider the pragmatic interpretation of relative clauses with and without the use of RPs (e.g. Cann et al., 2005; Collier, 1991; Erteschik-Shir, 1992; Galal, 2004) (see sections 3.3.2.2 and 3.3.2.3) . In this section, views on the interpretations of relative clauses with RPs versus gaps are discussed in light of the examples (126) and (127). Based on previous studies, there are three themes which might explain the use of the pronoun retention strategy. These themes are the concept, the RP for emphasis and the restrictive focus.

7.4.1 The Concept

Doron (1982) has stated that relative clauses that use the gap strategy have two interpretations; either by the 'concept' theme (i.e. there is no particular referent in the mind of the speaker), or it means that there is a particular referent that the speaker is talking about. On the other hand, relative clauses with the RP have only one interpretation, where the 'concept' theme is not allowed.

This interpretation cannot be applied to the Arabic sentences (126) and (127) above. One possible reason is the verb *ḏakar* 'mention' indicates a referential act with the RP or with the gap. On the other hand, the 'concept' interpretation might be used to explain the reason behind the absence of the RP in example (128).

128.	tağid	al-ʿabd	al-ṣāliḥ	allaḏī	taṭlub Ø
	you.find	the-worshiper	the-good	REL[3.SG.MAS]	you.seek Ø
	You will find the good worshipper that you seek.				
	[CA, 4: 125]				

In example (128), as can be understood from the context, the identity of the worshipper is not known for the one who seeks him. That person knows that there is a good worshipper, who is considered as the best on earth, but he does not know who that man is. He is looking for that person with these characteristics and it does not matter for him whether that person is Zaid or Ali. There is no particular worshipper is referred to in (128); therefore, the concept theme is found in this sentence which explains why the gap is chosen over the RP.

7.4.2 RPs for Emphasis

A number of previous studies have supported the emphatic function of the RP in SU position (Jassim, 2011), and in DO position (Givón, 1979) (see section 3.3.2.2).

Because the Arabic language is a null-subject language, one of the uses of the pronoun in the subject position is to emphasise the subject (Ouhalla, 1999; Ryding, 2005, p. 299) (see section 3.2.1.3). Cann et al (2005, p. 35) state that the RP, if used in SU relative clauses, will be 'phonologically stressed'. It is also indicated that the RP is used in a strong form in SU relative clauses in Lebanese Arabic (Aoun et al., 2001).

The argument that the RP is used for emphatic purposes in SU relative clauses is motivated by the repetition of the subject in the relative clause, where repetition is used to show emphasis (Hannouna, 2010). Subjects are usually implied in the form of the verb, if the relative clause has one, since Arabic verbs are inflected for number and gender, (see section 3.2.1.1 for more detail). Therefore, if the RP is used in the SU relative clause, it will be seen as a repetition of the subject of the relative clauses, as in example (96), which is repeated here as (129).

129.	al-faʾl	allaḏī	*huwa*	yurīdu	al-ẖair
	the-augury	REL(3.SG.MAS)	*it*	want(**it**)	the-good
	The augury which bodes the good things.				
	[CA, 2: 399]				

Using the 'emphasis' theme to show the difference between the two DO relative clauses (126) and (127) above could lead to the suggestion that in example (126), where the RP- *hu* 'it'- is focused, the head of the relative clause- *al-isnad* 'the reference'- is also focused because it is co referential with the RP. On the other hand, sentence (127) is neutral where there is no observable focus. However, in this study, the 'emphatic' interpretation is difficult to prove, especially in relative clauses other than SU relative clauses, for two reasons. First, repetition is not as clear in relative clauses other than SU as it is in SU relative clauses. For example, in DO relative

clauses, the RP is the only direct object inside the relative clause so the RP is not seen as repetition. Second, there is no marker for emphasis in Arabic. Moreover, to prove that the RP is phonologically stressed, spoken data is needed, which is not included in this study.

7.4.3 The Restrictive Focus

Erteschik-Shir (1992) states that the different interpretation which is related to the presence of the RP is a result of one of the RP's functions, which the author calls, "the restrictive focus". Using the RP in the relative clause is to restrict the focus of the listener to one referent rather than other referents which the listener knows about (see examples (53) and (54) in section 3.3.2.3).

This theme does not seem relevant to the case of the Arabic examples (125) and (126) mentioned above. Suggesting that the use of the RP (*-hu*) restricts the act of the mentioning (*ḏakar-*) to a particular referent (*-isnad*) would imply that the writer does not mention the other referents; this is not possible because there is only one author of the book. Thus, the 'restrictive focus' cannot account for examples (125) and (126) mentioned above.

7.5 The Use of the Pronoun Retention Strategy across Varieties of Arabic and genres

The results show that the use of RPs is different among the three varieties of Arabic. Whereas CA and IA show tendencies towards using RPs in relative clauses, MSA has a preference towards using the gap strategy, as is shown in Figure 35.

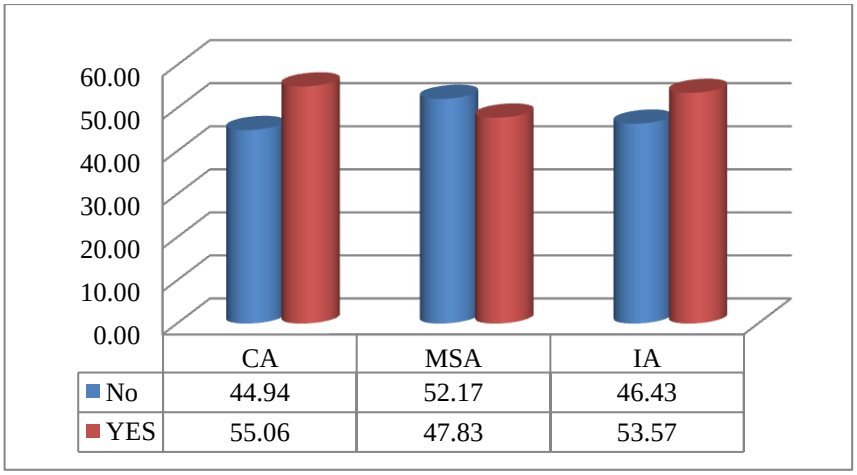


Figure 35: A comparison of the percentages of relative clauses that use RPs and relative clauses that use gaps across the three varieties of Arabic, where 'yes' indicates the use of RPs and 'no' indicates the use of gaps

The differences among varieties of Arabic in the use of the pronoun retention strategy might be attributed to the stylistic differences among the varieties. An example of stylistic changes is the use of *yuqāl la-hu* 'said to-it/him' in CA SU relative clauses, which is used 64 times (9.96% of all of relative clauses that use RPs), as in example (103), which is repeated here as (130).

130.	kān	fī	masǧidin	yuqāl	la-*hu*	masǧid	ṣaliḥ
	was	in	mosque	say(PASS)	to-*it*	mosque	Salih

He was in a mosque which is called Salih's mosque.
[CA, 4: 80]

yuqāl la-hu 'said to-it/him' is no longer used in MSA; it is substituted by *yusamā* 'called' where no RPs are needed. In IA, RPs are used consistently in all relative clauses except SU relative clauses (44.64%) and (3) OBL relative clauses.

A comparison between the two genres shows that while relative clauses have the tendency to use the pronoun retention strategy more in fiction, relative clauses in nonfiction texts have the tendency to use the gap strategy. However, the difference

between the percentage of relative clauses that use RPs and those which use gaps is very small in nonfiction texts, as is shown in Figure 36.

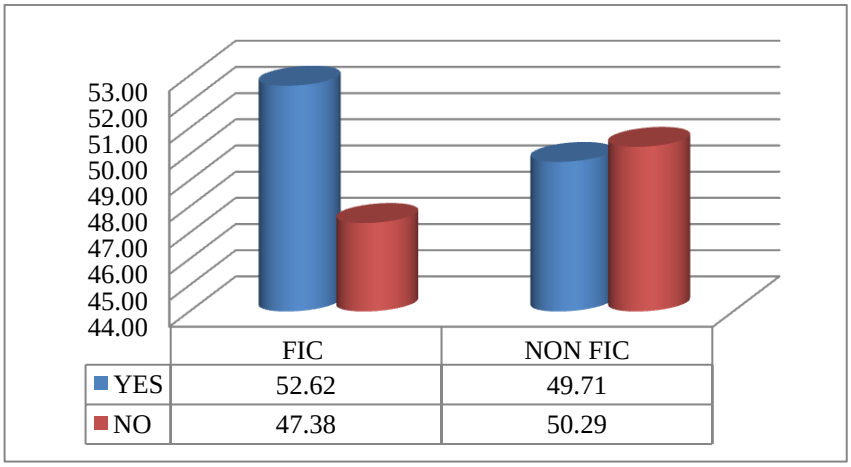


Figure 36: A comparison of the percentages of relative clauses that uses RPs and relative clauses that use gaps in fiction and nonfiction, where 'yes' indicates the use of RPs and 'no' indicates the use of gaps

The reason behind the difference in using the pronoun retention strategy in fiction and nonfiction texts is not clear, and might be worth studying further.

7.6 Summary

Results of this study show that the pronoun retention strategy is used in lower positions of the NPAH more than in the upper positions, which conforms to the NPAH's predictions. However, there are 128 violations to the claims that RPs are used only in the DO position and other positions down the NPAH, and RPs are not used in the SU position. These violations are distributed in different percentages among the three grammatical positions, SU, DO and OBL. One of the factors that determine the use of the pronoun retention strategy and interferes with the NPAH's predictions is the grammar of the Arabic language.

The grammar of the Arabic language, which prohibits preposition stranding, plays an important role in the use of RPs in relative clauses. In SU and DO relative clauses whose verbs are prepositional, pronouns must be retained because there is no preposition stranding in Arabic. Thus, regardless of the grammatical position of the relativised noun phrase, the pronoun retention strategy is the only strategy used in relativisation if the verb used in the relative clauses is a prepositional verb. The rule of the preposition stranding also governs the use of RPs in OBL since prepositions will not be used if RPs are absent.

The results of this study also show that there are certain factors that help in predicting the use of the pronoun retention strategy. The tendency towards using the pronoun retention strategy is found to be higher in indefinite relative clauses. This might also be related to the grammar of Arabic, where relative markers that are inflected for gender and number are not used in indefinite relative clauses. The results also show that counting several factors simultaneously such as the distance between the head and the relativised noun phrase, the complexity of the head, and the obligatoriness of the relativised noun phrase can predict the use of RPs more accurately.

The use of RPs differs across the three varieties of Arabic and the two genres. While no explanation was found for the difference between the two genres, the differences among varieties can be attributed to the stylistic changes. The differences among varieties and genres might be related to other factors mentioned earlier.

In this chapter, previous claims on the pragmatic interpretation of relative clauses have been discussed in light of the results of this study. All attempts to explain the use of RPs in relative clauses pragmatically treat the RP as an ordinary pronoun. The interpretations of relative clauses with or without RPs are based on different themes such as concept, focus, restrictive focus and the immediate context. It has been

found that one theme cannot adequately account for the interpretation of different relative clauses in different grammatical positions and in different contexts.

Chapter Eight: Conclusions

8.1 Introduction

This study has been conducted to examine the accessibility of different grammatical positions to relativisation across three varieties of Arabic: Classical Arabic, Modern Standard Arabic and Iraqi Arabic. Additionally, this study has attempted to investigate relativisation patterns in Arabic written fiction and nonfiction texts. With that goal, answers were sought regarding the following research questions:

1. Do the patterns of relativisation in Arabic follow the NPAH and/or the AH?
2. Is the distribution of relative clauses influenced by the variety of Arabic in which relative clauses are used?
3. Is the distribution of relative clauses influenced by the genre in which relative clauses are used?
4. Do characteristics of the heads of the relative clauses such as definiteness, animacy and the grammatical position of the head in the main clause affect the distribution of relative clauses?
5. How do syntactic and semantic factors affect the use of the pronoun retention strategy in Arabic relative clauses?

This chapter will present first in section 8.2 a summary of this study and how the methodology which has been used in this study has helped to answer the above questions. Then, conclusions on the general distribution of relative clauses and accessibility in relativisation in Arabic will be presented in section 8.3, which includes the effects of variety and genre on the distribution of relative clauses in sections 8.3.1 and 8.3.2, respectively, and the role of the characteristics of the head in the distribution

of relative clauses in section 8.3.3. Next, the conclusions drawn on the use of the pronoun retention strategy will be presented in 8.4. Finally, the implications of this study and recommendations for future studies are presented in 8.5 and 8.6, respectively.

8.2 Study Design

As is shown in the methodology chapter, the data of this study have been collected from fiction and nonfiction texts that are written in three varieties of the Arabic language, CA, MSA and IA. In analysing the data, different factors have been considered; these factors are definiteness, animacy and the grammatical position of the head in the main clause. In this study, multi-level Poisson regression analysis has been used since this statistical method can account for multiple factors when predicting counts (Baayen, 2008; Keune et al., 2005). Multi-level Poisson regression analysis enhances the reliability of the results because it considers any differences which might take place among individual texts.

8.3 The General Distribution of Relative Clauses in Arabic Written Texts and Accessibility to Relativisation

As shown in the Chapter Five, the overall distribution of relative clauses in Arabic written texts supports the NPAH except for the IO position, since IO relative clause occur rarely, only twice, in this study. The AH, however, is more strongly supported than the Subject Primacy Hypothesis because ABS relative clauses are found significantly more than SU relative clauses in the overall distribution of relative clauses in the corpus. However, the AH is not always supported; while the difference between ABS relative clauses and SU relative clauses is significant in CA, it is not significant in MSA and IA. Thus, the results seem to suggest a new hierarchy while including ABS

relative clauses as a category and including the two grammatical positions IO and OCOMP, under one category "OTHERS", as both types of relative clauses are rarely found. This hierarchy is as follows: ABS>TSU,OBL> GEN>OTHERS, where (>) indicates more accessible and (,) indicates that there is no difference between the two positions.

The discourse functions of relative clauses, such as characterising and anchoring, help in determining how frequently these relative clauses are used in written corpora. However, the results have shown that the frequency of OBL relative clauses is not always related to their functions. The other noun phrase inside the OBL relative clause serves as a better anchor than DOs in TSUs relative clauses, yet OBL relative clauses, except in CA, are not found more frequent than TSU relative clauses. This indicates that the correlation between the frequency of relative clauses and their description and anchoring functions is found to be dominant only in the three grammatical positions (ISU, TSU, DO) and is not so influential in other grammatical positions. It follows that the distribution of relative clauses in written texts is not only motivated by the functions of relative clauses in the texts. That is, the functions which relative clauses perform in the texts cannot adequately predict the frequency order of these relative clauses in the text.

As a summary, the accessibility in relativisation cannot be determined by either the grammatical positions of the relativised noun phrases or the functions of relative clauses in discourse alone. In fact, accessibility in relativisation can be motivated by both of these factors, the discourse functions and the grammatical positions in the relative clause, among other factors. These factors have been examined in this study under two categories: (1) construction-external factors; (2) construction-internal factors. The conclusions on construction-external factors, variety and genre, are presented in

8.3.1 and 8.3.2, respectively. The effects of construction-internal factors will be presented in section 8.3.3.

8.3.1 Relativisation across Varieties

Each variety of Arabic has revealed a different pattern of relativisation. The distribution of MSA is closer to IA than to CA. As shown in section 5.6, IA is different from MSA in two positions only and both of these varieties differ from CA in four positions. The overall distribution of the whole number of relative clauses in Arabic written texts does not match with any of the distributions revealed by the varieties. The AH is reflected in the overall distribution of relative clauses as well as in CA, but not in MSA and IA. The hierarchy suggested above, ABS> TSU,OBL> GEN>OTHERS is reflected in MSA and IA but not in CA.

These results seem to suggest two conclusions; first, the distribution of relative clauses is influenced by the variety in which it occurs. Second, the Arabic language cannot be represented by any single variety; that is, a sample of CA relative clauses is not enough for studying relativisation in Arabic. In general, therefore, these results suggest that it is important to consider different varieties of the same language in deciding accessibility to relativisation in that language. Results also bring up a question of what accessibility in a diglossic situation is. In other words, does an individual who speaks both MSA and IA have two different accessibility hierarchies in his mind? The results show that the difference between CA and the other varieties makes CA appears as another language rather than a variety of the Arabic language. This suggests that speakers of the these varieties have two different accessibility hierarchies, one

represents MSA and IA and the other represents CA, where they switch from one to another according to the variety they use.

8.3.2 Relativisation in Fiction and Nonfiction

Relative clauses in fiction have displayed an order that is slightly different from the order of relative clauses in nonfiction texts, yet, the AH is reflected in the distributions of both genres as the ABS relative clauses outrank the SU relative clauses. In general, therefore, it seems that the differences between the distribution of relative clauses in fiction and nonfiction texts are less than the differences among the three varieties. The distribution of relative clauses in nonfiction texts is identical to the overall distribution of relative clauses in Arabic texts when including ISU and DO relative clauses under the ABS category, as in the hierarchy suggested earlier: ABS> TSU, OBL> GEN>OTHERS. On the other hand, in fiction texts, TSU relative clauses are found significantly more than OBL relative clauses.

8.3.3 The Effect of the Construction-Internal Factors on the Frequency of Relative Clauses.

As explained before, there are three construction-internal factors that have been considered in analysing the data of this study. The first of these factors is definiteness. The results have shown that definite relative clauses have a higher frequency than indefinite relative clauses in Arabic written texts. This result shows a consistency across the three varieties and in the two genres. It follows that there is a tendency to use definite relative clauses more than indefinite relative clauses, and this tendency is not influenced by the genre or the variety in which the text is written. The distribution of definite relative clauses is slightly different from the distribution of indefinite relative

clauses. The distributions of definite and indefinite relative clauses differ across the three varieties of Arabic. These results, thus, suggest the distribution of relative clauses in written texts is affected by varieties as well as the definiteness of the head.

The second factor that is considered in this study is the grammatical position of the head in the main clause. Results show that relative clauses in certain grammatical positions have higher frequencies than others. Moreover, noun phrases in some grammatical positions, OBL, DO and SU, show a higher tendency to be identified by ISU relative clauses; others such as GEN, SU(TOPIC) and SUCOM prefer to be identified by DO, and relative clauses in one grammatical position 'OTHERS' show a tendency to be identified by TSU. This suggests two points: First, the position of the head of relative clauses in the main clause contributes to the frequency of relative clauses. Second, heads in all grammatical positions prefer to be identified by relative clauses in grammatical positions higher in the NPAH, which manifests the accessibility claim that relative clauses in higher grammatical positions in the NPAH such as TSU, ISU and DO are more accessible than relative clauses in lower positions such as OBL and GEN.

The third factor considered in this study is the animacy of the head. The role of animacy in the distribution of relative clauses is quite important since relative clauses that refer to animate referents reflect the SPH whereas relative clauses that refer to inanimate referents follow the AH. The role of animacy does not only affect relative clauses in the higher positions, TSU, ISU, and DO, but it also affects the distribution of relative clauses in OBL and GEN positions. This is found consistently across the three varieties and the two genres.

As has been shown in Chapters Five and Six, there is no positive relation between definiteness and the animacy of the head in the structures of relative clauses. This suggests that definiteness does not depend on the animacy of the head noun phrase in the relative clause. On the other hand, the results suggest that definiteness is related to the grammatical position of the relativised noun phrase and animacy.

It can be concluded that definiteness, animacy and the grammatical position of the head all contribute to the frequency of relative clauses. The animacy of the head affects the frequency as well as the distribution of relative clauses in the written texts. Thus, the evidence from this study suggests that the role of animacy in predicting the order of the frequency of relative clauses in written texts is more influential than the roles of the other two factors, which emphasises the necessity to consider animacy in similar future studies.

Taken together, the results of this study suggest that the accessibility to relativisation in the Arabic language is motivated by a number of construction-external and construction-internal factors in addition to the grammatical position of the head in the relative clause, the NPAH, and the discourse functions of relative clauses.

8.4 The Use of the Pronoun Retention Strategy in Arabic Relative Clauses

In this study, an attempt has been made to explain the basis on which the pronoun retention strategy must be used, and where it can be optional. The use of RPs has been studied in relation to different theories such as the NPAH (Keenan & Comrie, 1977) and the Accessibility Theory (Ariel, 1990, 1999) . Other syntactic and semantic factors such as the distance between the head and the relativised noun phrase,

definiteness and the animacy of the head have also been considered to explain the use of the pronoun retention strategy. As has been shown in Chapter Seven, descriptive analysis is used to explain the results instead of multi-level Poisson regression analysis because of the small numbers of relative clauses under consideration. The results show that the RP is used in all of the grammatical positions which are tested in this study: TSU, ISU, DO, IO, OBL, OCOMP. On the other hand, the gap strategy is used in all the mentioned positions except IO, GEN, OCOMP. Accordingly, the following conclusions have been suggested:

- The RP is used mainly to avoid syntactic and semantic ambiguity. The absence of the RPs in some relative clauses turns them into ambiguous sentences, as the reader cannot decide whether they are SU or DO relative clauses; in such cases, the RP is actually required.
- RPs are used either optionally, or as a 'last option' to save sentences that are ungrammatical without the use of the pronoun retention strategy. The grammaticality of the Arabic relative clauses is mainly governed by the rules of the Arabic language such as the prohibition of preposition stranding. According to this rule, the RP is required whenever a preposition is used regardless of the grammatical position that is relativised.
- The RP is preferred in the least accessible relative clauses. Therefore, it is used in grammatical positions down the NPAH such as OBL and GEN more than in DO and SU positions, where the RP is rarely used. It is also used in indefinite more than definite relative clauses and in relative clauses where the distance between the head and the relativised noun phrase is large.

- Considering more than one factor at the same time, as is suggested by the Accessibility Theory, has provided better predictions on the use of the RPs.
- Pragmatic factors could explain the reason why the RP is preferred over the gap in sentences where both strategies are acceptable. The use of the RP might be explained through one of the following interpretations: the focus, the concept, the restrictive focus and the immediate context; however, the results suggest that there is no single pragmatic interpretation that can account for all the different instances where the pronoun retention is optionally used.

8.5 Limitations and Recommendations for Future Research

As discussed in the methodology chapter, the differences among the texts have not been considered in the data collection. That is, some texts reveal a very high number of relative clauses in comparison to the other texts. There have been no attempts to limit the collection to a specific number of relative clauses in all the books. However, this was partially accounted for with the method of statistical analysis. It is necessary that difference among texts be controlled in future studies, either in the data collection process or in the analysis of the data.

One of the limitations of this study is that this study does not consider nonfiction Iraqi texts. That was because a library search did not reveal nonfiction texts written in the Iraqi dialect. This situation might change in the future, and then it will be interesting to include nonfiction Iraqi texts in future studies. Similarly, it is difficult to find fiction texts written in CA, therefore, the fiction CA texts included in this study are not first-person narratives like the MSA fiction texts.

CA is older than MSA, yet this study is not historical in the sense that it has not examined the language change over time. It will be interesting for future studies to deal with this chronological dimension and study the changes that relative clauses undergo. It will be also interesting to test whether the changes happened over time follow the NPAH's predictions or not.

The results of this study yield a high frequency for OBL relative clauses in CA texts. The popularity of OBL for relativisation in CA, but not in MSA and IA is worth further exploring. The data of this study include written texts of IA, where only the conversations among the characters are considered. In future research, it is worth including spoken data recorded from everyday conversations to be compared with written data to examine the distribution of relative clauses, definiteness, animacy, and so forth.

The effect of diglossia on the use of relative clauses in Arabic texts in this study is tested based on the differences among the three varieties, although the texts are written by different writers. In future studies, it is recommended to include texts written in an Arabic dialect one time and in MSA in another by the same writer as this could improve the validity of the results.

In this study, first-person narratives only are included in MSA fiction and only history books are included in nonfiction. Other types of text could also be added to the future research. For example, including third person narratives is a good way to compare the distribution of SU and DO relative clauses in first- and third person narratives. Similarly, including other types of nonfiction such as textbooks or biographies provides the chance to examine whether all nonfiction texts reveal the same distribution of relative clauses.

Animacy is found to be an important factor that affects the distribution of relative clauses, yet only the animacy of the head has been considered in this study. Previous studies considered the animacy of the other noun phrase in the relative clause such as the SU in the DO relative clause and DO in SU relative clauses. Therefore, considering the animacy of the other noun phrases in the relative clause could reveal important results that may enhance studies on Arabic relativisation in general and Arabic corpus studies in particular.

For the use of the pronoun retention strategy in Arabic relative clauses, it would be interesting to include non-restrictive relative clauses, since it is one of the factors which are considered by Ariel (1999), but could not be tested in this study because of the limitation of our data to restrictive relative clauses only. It would also be interesting to add more factors to those included under the Accessibility Theory, especially the factors that have been found to be influential in this study, such as the definiteness of the head.

References

Abdelghany, H. (2010). *Prosodic phrasing and modifier attachment in standard Arabic sentence processing.* Unpublished doctoral dissertation, City University of New York New York.

Al-anbari, A. A. (1957). *Asrar Al-3arabia 'Secrets of Arabic'*. Damascus: Al-majma3 Al-3ilmi Al-3arabi.

Al-bazi, M. A. (1983). *Scopal properties of the Arabic relative clause.* Unpublished doctoral dissertation, University of Washington, Washington.

Al-dahdah, A. F. (2007). *Sharh alfiat Ibn Malik 'Explanation of alfiat Ibn Malik'*. Riyadh: Obeikan Bookshop.

Al-fayyuumi, A. (1987). *Al-miṣbaaḥ al-muniir: Mu3jam 3arabi-3arabi 'The Enlightening lamp: Arabic- Arabic dictionary'* Beirut: Maktabat Lubnaan.

Al-malyaani, M. (1979). *Mu3jam al-af3aal al-muta3addiyah bi-ḥarf al-jar ' Dictionary of transitive prepositional verbs'*. Beirut: Daar Al-3ilm lil-Malayiin.

Al-muradi, I. U. Q. (2001). *Tawdhiih Al-maqaasid wa al-masaalik fi sharh alfiat Ibn Malik ' Clarification of purposes and tracks in the explanation of alfiat Ibn Malik'* (Vol. 2). Cairo: Dar Alfikr Al-3arabi.

Al-raazi, M. (1988). *Mukhtar al-sihahh 'Mukhtar al-sihah dictionary'*. Beirut: Maktabat Lubnaan.

Al-yasiri, S. (1972). *Ruba'iyaat Abu Gaaṭi' 'Abu Gaaṭi''s Quadruplet'*. Baghdad: Maṭba3at Al-sha3b.

Al-zajaji, A. A. (1979). *Al-idaaḥ fi al-naḥu 'Illustrations in grammars'*. Beirut: Dar Al-nafa'is.

Al-zamakhshari, M. (1984). *Asaas al-balaaghah 'The basis of eloquence'* Beirut: Daar Bayrut lil-Tib'ah wa-al-Nashr.

Allen, R. (1998). *The Arabic literary heritage: the development of its genres and criticism*. Cambridge: Cambridge University Press Cambridge.

Allen, S. E. M., & Schroder, H. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In J. W. Du Bois, L. E. Kumpf & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (pp. 301-338). Amestrdam: John Benjamins Publishing Company

Altoma, S. (1969). *The problem of diglossia in Arabic: A comparative study of classical and Iraqi Arabic*. Cambridge: Harvard University Press.

Aoun, J., Choueiri, L., & Hornstein, N. (2001). Resumption, movement, and derivational economy. *Linguistic Inquiry, 32*(3), 371-403.

Aoun, J. E., Choueiri, L., & Benmamoun, E. (2010). *The syntax of Arabic*. Cambridge: Cambridge University Press.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, M. (1999). Cognitive universals and linguistic conventions: The case of resumptive pronouns. *Studies in Language, 23*(2), 217-269.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Badawi, E., Carter, M. G., & Gully, A. (2004). *Modern written Arabic: A comprehensive grammar*. London: Routledge.

Bahloul, M., & Harbert, W. (1993). Agreement asymmetries in Arabic. In J. Mead (Ed.), *Proceedings of the eleventh west coast conference on formal linguistics* (pp. 15-31). Stanford: Centre for the Study of Language and Information.

Bashir, S. M. (1982). *A functional approach to Arabic relative clauses with implications for foreign language teaching.* Unpublished Ed.D dissertation, Columbia University Teacher College, New York.

Bassiouney, R. (2006). *Functions of code switching in Egypt: Evidence from monologues*. Leiden: Brill.

Bassiouney, R. (2009a). *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.

Bassiouney, R. (2009b). The variety of housewives and cockroaches: Examining code-choice in advertisements in Egypt. In E. Al-Wer & R. D. Jong (Eds.), *Arabic dialectology: In honour of Clive Holes on the occasion of his sixtieth birthday* (pp. 273-283). Leiden: Brill.

Bickel, B. (2010). Grammatical relations typology. In J. J. Song (Ed.), *The Oxford handbook of language typology* (pp. 399-444). Oxford: Oxford University Press.

Bošković, Ž. (2009). On relativization strategies and resumptive pronouns. In G. Zypatow, U. Junghanns, D. Lenertova & P. Biskup (Eds.), *Studies in formal Slavic phonology, morphology, syntax, semantics and information structure. Proceedings of FDSL 7, Leipzig* (pp. 79-93). Bern: Peter Lang International Academic Publishers.

Brustad, K. (2000). *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington D.C.: Georgetown University Press.

Cann, R., Kaplan, T., & Kempson, R. (2005). Data at the grammar-pragmatics interface: The case of resumptive pronouns in English. *Lingua, 115*(11), 1551-1577.

Chen, Z., Li, Q., Kuo, K.-L., & Vasishth, S. (2010). *Processing Chinese relative clauses: Evidence for the universal subject preference*. Cornell University, Department of Linguistics. New York.

Cinque, G. (1981). On Keenan and Comrie's primary relativization constraint. *Linguistic Inquiry, 12*(2), 293-308.

Clancy, P. M., Lee, H., & Zoh, M.-H. (1986). Processing strategies in the acquisition of relative clauses: Universal principles and language-specific realizations. *Cognition, 24*(3), 225-262.

Cole, P. (1976). The interface of theory and description: Notes on modern Hebrew relativization. *Language, 52*(3), 563-583.

Cole, P., Harbert, W., Sridhar, S., Hashimoto, S., Nelson, C., & Smietana, D. (1977). Noun phrase accessibility and island constraints. In P. Cole & J. M. Sadock (Eds.), *Syntax and semantics* (pp. 27-46). London: Academic Press.

Cole, P., & Hermon, G. (2005). Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics, 14*(1), 59-88.

Collier, M. (1991). The relative clause and the verb in middle Egyptian. *The Journal of Egyptian Archaeology*(77), 23-42.

Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology* (2nd ed.). Chicago: University of Chicago press.

Comrie, B., & Kuteva, T. (2005). Relativization strategies. In M. Haspelmath, M. Dryer, D. Gil & B. Comrie (Eds.), *The world atlas of language structures* (pp. 494- 499). Oxford: Oxford University Press.

Comrie, B., & Kuteva, T. (2011). Relativization strategies M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online* Retrieved from http://wals.info/supplement/8

Croft, W. (2003). *Typology and universals* (2nd ed.). Cambridge: Cambridge University Press.

Croteau, K. (1995). Second language acquisition of relative clause structures by learners of Italian. In F. R. Eckman, D. Highland, P. W. Lee, J. Mileham & R. R. Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 115–128). Mahwah: Lawrence Erlbaum Associates.

Cuvalay-Haak, M. (1997). *The verb in literary and colloquial Arabic*. Berlin: Walter de Gruyter.

Danks, W. (2011). *The Arabic verb: Form and meaning in the vowel-lengthening patterns*. Amsterdam: John Benjamins Pubishing Company.

Dekeyser, X. (1984). Relativizers in early modern English: A dynamic quantitative study. In J. Fisiak (Ed.), *Historical syntax* (pp. 61-87). Berlin: Mouton De Gruyter.

Diesing, M. (1992). *Indefinites*. Cambridge, Mass: MIT Press.

Diessel, H. (2009). On the role of frequency and similarity in the acquisition of subject and non-subject relative clauses. In T. Givón & M. Shibatani (Eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution* (pp. 251-276). Amsterdam: John Benjamins Publishing Company.

Doron, E. (1982). On the syntax and semantics of resumptive pronouns. *Texas Linguistic Forum Austin*(19), 1-48.

Doron, E. (2011). On the syntax and semantics of resumptive pronouns. In A. Rouveret (Ed.), *Resumptive pronouns at the interfaces* (pp. 289-318). Amsterdam: John Benjamins Publishing Company.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*(04), 431-469.

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language, 63*(4), 805-855.

Eckman, F., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*(1), 1-20.

Eid, M. (1982). The non-randomness of diglossic variation in Arabic. *Glossa, 16*(1), 54-84.

Eid, M. (1983). On the communicative function of subject pronouns in Arabic. *Journal of Linguistics, 19*(02), 287-303.

Eid, M. (1988). Principles for Code-Switching between Standard and Egyptian Arabic. *al-Arabiyya, 21*(1), 51-79.

Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry, 22*(1), 1-25.

Ennaji, M. (2005). *Multilingualism, cultural identity, and education in Morocco*. New York: Springer.

Erteschik-Shir, N. (1992). Resumptive pronouns in islands. In H. Goodluck & M. Rochemont (Eds.), *Island constraints: Theory, acquisition, and processing* (pp. 89-108). Dordrecht: Kluwer Academic Publishers.

Eviatar, Z., & Ibrahim, R. (2012). Multilingualism among Israeli Arabs, and the neuropsychology of reading in different languages. In M. Leikin, M. Schwartz & Y. Tobin (Eds.), *Current Issues in Bilingualism: Cognitive and socio-linguistic perspectives* (pp. 57-74). Dordrecht: Springer.

Ferguson, C. (1959). Diglossia. *Word, 15*(2), 325-340.

Fernandez-Duque, D. (2009). Cognitive and neural underpinnings of syntactic complexity. In T. Givón & M. Shibatani (Eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution*. Amsterdam: John Benjamins Publishing Company.

Fleischer, J. (2004). A typology of relative clauses in German dialects. In B. Kortmann (Ed.), *Dialectology meets typology: Dialect grammar from a cross-linguistic perspective* (pp. 211–243). Berlin: Mouton de Gruyter.

Fox, B. A. (1987). The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis? *Language, 63*(4), 856-870.

Fox, B. A., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language, 66*(2), 297-316.

Galal, M. M. (2004). *A minimalist approach to relative clauses in modern standard Arabic.* Unpublished doctoral dissertation, University of Kansas, Kansas.

Gary, J., & Gamal-Eldin, S. (1981). *Cairene Egyptian colloquial Arabic*. Kent: Croom Helm.

Gas, S. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*(2), 327-344.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Givón, T. (1993). *English grammar: A function-based introduction* (Vol. 2). Amestrdam: John Benjamins Publishing Company.

Godfrey, J. J., Holliman, E. C., & McDaniel, J. (1992). *SWITCHBOARD: Telephone speech corpus for research and development*.

Gómez, C. S. (2006). *Relativization in early English (950-1250): The position of relative clauses* (Vol. 49). Bern: Peter Lang Pub Incorporated.

Gordon, P., & Hendrick, R. (2005). Relativization, ergativity, and corpus frequency. *Linguistic Inquiry, 36*(3), 456-463.

Hamdallah, R., & Tushyeh, H. (1998). A contrastive analysis of English and Arabic in relativization. *Papers and Studies in Contrastive Linguistics, 34*, 141-152.

Hannouna, Y. H. (2010). Assessment of translating recurrence in selected texts from the Qur'an. *Journal of Translation and Technical Communication Research Trans-kom, 3*(1), 85-113.

Hardy, D., & Milton, K. (1994). The distribution and function of relative clauses in literature. In L. Bouton & Y. Kachru (Eds.), *Pragmatics and Language Learning. Monograph Series* (Vol. 5, pp. 247-265). University of Illinois: Urbana-Champaign.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, J. A. (1999a). Processing complexity and filler-gap dependencies across grammars. *Language, 75*(2), 244-285.

Hawkins, J. A. (1999b). The relative order of prepositional phrases in English: Going beyond manner–place–time. *Language Variation and Change, 11*(03), 231-266.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hawkins, R. (2007). The noun phrase accessibility hierarchy: Lame duck or dead duck in theories of SLA? *Studies in Second Language Acquisition, 29*(02), 345-349.

Heestand, D., Xiang, M., & Polinsky, M. (2011). Resumption still does not rescue islands. *Linguistic Inquiry, 42*(1), 138-152.

Herrmann, T. (2005). Relative clauses in English dialects of the British Isles. In L. W. Pietsch, Susanne; Kortmann, Bernd (Ed.), *A comparative grammar of British English dialects: agreement, gender, relative clauses* (pp. 21–124): Mouton de Gruyter, Berlin.

Hogbin, E., & Song, J. (2007). The accessibility hierarchy in relativisation: The case of eighteenth- and twentieth-century written English narrative. *SKY Journal of Linguistics, 20*, 203-233.

Holes, C. (1990). *Gulf Arabic*. London: Routledge.
Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, D.C.: Georgetown University Press.
Hyltenstam, K. (1984). The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Andersen (Ed.), *Second languages: A cross-linguistic perspective* (pp. 39-58). Rowley, Mass: Newbury House.
Ioup, G., & Kruse, A. (1977). *Interference versus structural complexity as a predictor of second language relative clause acquisition*. Paper presented at the Los Angeles Second Language Research Forum
Ishizuka, T. (2005). Processing relative clauses in Japanese. *Papers in Psycholinguistics, 2*, 135-157.
Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning, 53*(2), 285-323.
Jaggar, P. J. (2001). *Hausa*. Amsterdam: John Benjamins Publishing Company.
Janet, C., & Watson, E. (1993). *A Syntax of Ṣanʿānī Arabic*. Wiesbaden: Otto Harrassowitz Verlag.
Jassim, Q. H. (2011). *Relative clauses in Iraqi Arabic and the status of resumptive pronouns.* Unpublished master's thesis, Universitat Autònoma de Barcelona, Barcelona.
Jayyusi, S. K. (2010). *Classical Arabic stories: An anthology*. New York: Columbia University Press.
Jensen, H. B. (1999). *Reduced relative clauses used in eighteenth-century diaries and letters: A sociohistorical perspective.* Unpublished doctoral dissertation, Indiana University, Bloomington.
Jeon, K., & Kim, H. (2007). Development of relativization in Korean as a foreign language: The noun phrase accessibility hierarchy in head-internal and head-external relative clauses. *Studies in Second Language Acquisition, 29*(02), 253-276.
Johnson, D. E. (1974). *On the role of grammatical relations in linguistic theory.* Paper presented at the Chicago linguistic society, Chicago
Johnson, D. E. (1977). On Keenan's definition of "subject of". *Linguistic Inquiry, 8*(4), 673-692.
Joseph, B. (1980). Recovery of information in relative clauses: Evidence from Greek and Hebrew. *Journal of Linguistics, 16*(02), 237-244.
Kanno, K. (2007). Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition, 29*(2), 197-218.
Katzner, K. (2002). *The languages of the world*. London: Routledge.
Kaye, A. S. (2009). Arabic. In B. Comrie (Ed.), *The world's major languages* (2nd ed.). Abingdon: Routledge.
Keenan, E. (1975). Variation in universal grammar. In R. W. Fasold & R. W. Shuy (Eds.), *Analysing variation in language* (pp. 136-148). Washington DC: Georgetown University Press.
Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*(1), 63-99.
Keenan, E., & Comrie, B. (1979). Data on the noun phrase accessibility hierarchy. *Language, 55*(2), 333-351.

Keenan, E., & Hawkins, S. (1987). The psychological validity of the accessibility hierarchy. In E. Keenan (Ed.), *Universal grammar* (pp. 60-85). New Hampshire: Croom Helm.

Keune, K., Ernestus, M., Van Hout, R., & Baayen, R. (2005). Social, geographical, and register variation in Dutch: From written 'Mogelijk' to spoken 'Mok'. *Corpus Linguistics and Linguistic Theory, 1,* 183-223.

Kortmann, B. (2004). *Dialectology meets typology: Dialect grammar from a cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Kortmann, B., Herrmann, T., Pietsch, L., & Wagner, S. (2005). *Agreement, gender, relative clauses*. Berlin: De Gruyter Mouton.

Kortmann, B., & Wagner, S. (2005). FRED–The Freiburg English dialect corpus. In B. Kortmann (Ed.), *A comparative grammar of british English dialects: Agreement, gender, relative clauses* (pp. 1-20). Berlin: Mouton de Gruyter.

Kroch, A. S. (1981). On the role of resumptive pronouns in amnestying island constraint violations. In R. A. Hendrick, C. S. Masek & M. F. Miller (Eds.), *Papers from the 17th Regional Meeting, Chicago Linguistic Society* (pp. 125-135). Chicago: University of Chicago.

Kucera, H., & Francis, W. N. (1967). *The Brown University standard corpus of present day American English*. Rhode Island: The Brown Universiy.

Kuno, S. (1974). The position of relative clauses and conjunctions. *Linguistic Inquiry, 5*(1), 117-136.

Kwon, N., Gordon, P. C., Lee, Y., Kluender, R., & Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of prenominal relative clauses in Korean. *Language, 86*(3), 546-582.

Labelle, M. (1990). Predication, Wh-movement, and the development of relative clauses. *Language Acquisition, 1*(1), 95.

Lawal, A. S. (1987). Yoruba relativisation and the continuous segment principle. *Studies in African linguistics, 18*(1), 67-79.

Lin, C.-J., & Bever, T. G. (2006). Subject preference in the processing of relative clauses in Chinese. In D. Baumer, D. Montero & M. Scanlon (Eds.), *Proceedings of the 25th west coast conference on formal linguistics* (pp. 254-260). Somerville, MA: Cascadilla Proceedings Project.

Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (Vol. 2). East Sussex: Psychology Press.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*(1), 50-68.

Malchukov, A., Haspelmath, M., & Comrie, B. (2010). Ditransitive constructions: A typological overview. In A. Malchukov, M. Haspelmath & B. Comrie (Eds.), *Studies in ditransitive constructions: A comparative handbook* (pp. 1-64). Berlin: Walter de Gruyter.

Matsumoto, Y. (1990). The role of pragmatics in Japanese relative clause constructions. *Lingua, 82*(2-3), 111-129.

Maxwell, D. N. (1979). Strategies of relativization and NP accessibility. *Language, 55*(2), 352-371.

McCloskey, J. (1979). *Transformational syntax and model theoretic semantics: A case study in modern Irish*. Dordrecht: D. Reidel

Moutaouakil, A. (1989). *Pragmatic functions in a functional grammar of Arabic*. Dordrecht: Foris Publications Holland.

O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition, 25*(03), 433-448.

Ouhalla, J. (1999). Focus and Arabic clefts. In G. T. Rebuschi, Laurice (Ed.), *The grammar of focus* (pp. 335-359). Amsterdam: John Benjamins Publishing Company.

Ozeki, H. (2011). The acquisition of relative clauses in Japanese. In E. J. Kidd (Ed.), *The acquisition of relative clauses: Processing, typology and function* (pp. 173-194). Amsterdam: John Benjamins Publishing Company.

Ozeki, H., & Shirai, Y. (2007). Does the noun phrase accessibility hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition, 29*(02), 169-196.

Palmer, F. R. (1994). *Grammatical roles and relations*. Cambridge: Cambridge University Press.

Pashova, T. (2002). Classical Arabic and modern standard Arabic - Is there a process of change? In S. Leder, H. Kilpatrick, B. Martel-Thoumian & H. Schönig (Eds.), *Studies in Arabic and Islam* (Vol. 108, pp. 379-396). Leuven: Peeters Publishers.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition, 8*(01), 38-55.

Pérez-Leroux, & Teresa, A. (1995). Resumptives in the acquisition of relative clauses. *Language Acquisition, 4*(1-2), 105-138.

Perlmutter, D. M. (1980). Relational grammar. In J. R. Wirth & E. A. Moravcsik (Eds.), *Current approaches to syntax* (pp. 195-229). New York: Academic Press.

Phoocharoensil, S. (2010). Language universals and learning strategies in L2 relative clause acquisition. *International Journal of Learning, 17*(2), 95-108.

Polinsky, M., Gallo, C. G., Graff, P., & Kravtchenko, E. (2012). Subject preference and ergativity. *Lingua, 122*(3), 267-277.

Prideaux, G. D., & Baker, W. J. (1986). *Strategies and structures: The processing of relative clauses* (Vol. 46). Amsterdam: John Benjamins Publishing Company.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics*. New York: Academic Press.

Prince, E. F. (1990). Syntax and discourse: A look at resumptive pronouns. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society, 16*, 482-497.

Qabawa, F. A. (1989). *I3rab al-jumal wa ashbah al-jumal ' Analysis of sentences and semi sentences'* (5th ed.). Halab: Dar Al-itifaq Al-3arabia.

Reali, F., & Christiansen, M. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language, 57*(1), 1-23.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*(3), 348-379.

Romaine, S. (1980). The relative clause marker in Scots English: Diffusion, complexity, and style as dimensions of syntactic change. *Language in Society, 9*(02), 221-247.

Ross, J. R. (1967). *Constraints on variables in syntax.* PhD Dissertation, Massachusetts Institute of Technology, Cambridge.

Ryding, K. (1977). *Semantic and syntactic aspects of Arabic prepositions.* Unpublished doctoral dissertation, Georgetown University, Georgetown.

Ryding, K. (2005). *A reference grammar of modern standard Arabic*. Cambridge: Cambridge University Press.

Ryding, K. (2011). Arabic datives, ditransitives, and the preposition *li-*. In B. Orfali (Ed.), *In the shadow of Arabic: The centrality of language to Arabic culture* (Vol. 63, pp. 283-299). Leiden, The Netherlands: Brill.

Saiegh-Haddad, E. (2012). Literacy reflexes of Arabic diglossia. In M. Leikin, M. Schwartz & Y. Tobin (Eds.), *Current issues in bilingualism: Cognitive and socio-linguistic perspectives* (pp. 43-55). Dordrecht: Springer.

Salzmann, M. (2006). Resumptive pronouns and matching effects in Zurich German relative clauses as distributed deletion. *Leiden Papers in Linguistics, 3*(1), 17-50.

Sells, P. (1984). Syntax and semantics of resumptive pronouns. *Linguistic Review, 4*(3), 261-267.

Sells, P. (1987). Binding resumptive pronouns. *Linguistics and Philosophy, 10*(3), 261-298.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior, 13*(3), 272-281.

Shlonsky, U. (1992). Resumptive pronouns as a last resort. *Linguistic Inquiry, 23*(3), 443-468.

Sibawayh, I. A. (1988). *Al-kitaab 'The book'*. Cairo: Maktabat Al-khanchi.

Sigurd, B. (1989). A supplementary relativization hierarchy based on the complexity of the relative phrase. *Studia linguistica, 43*(1), 33-46.

Slobin, D. I. (1971). Developmental psycholinguistics. In W. O. Dingwall (Ed.), *A survey of linguistic science.* (pp. 298-412). College Park: University of Maryland (Linguistics program).

Song, J. J. (1991). Korean relative clause constructions: Conspiracy and pragmatics. *Australian Journal of Linguistics, 11*(2), 195-220.

Song, J. J. (2001). *Linguistic typology: Morphology and syntax*. Harlow: Pearson Education

Song, J. J. (2003). Resumptive genitive pronouns in Korean relative clauses: Distribution and explanation. *SKY Journal of Linguistics*(16), 139-160.

Suaieh, S. I. (1980). *Aspect of Arabic relative clauses: A study of the structure of relative clauses in modern written Arabic.* Doctoral dissertation, Indiana University, Bloomington.

Suárez, C. (2012). The consolidation of þat as an invariable relativizer in the history of English. *Nordic Journal of English Studies, 11*(1), 79-107.

Suleiman, S. M. (1985). *Jordanian arabic between diglossia and bilingualism: linguistic analysis*. Amsterdam: John Benjamins Publishing.

Suñer, M. (1998). Resumptive restrictive relatives: A crosslinguistic perspective. *Language, 74*(2), 335-364.

Tarallo, F. (1986). Functional and structural properties in a variable syntax. In D. Sankoff (Ed.), *Diversity and diachrony*. Amsterdam: John Benjamins Publishing Company.

Tarallo, F., & Myhill, J. (1983). Interference and natural language processing in second language acquisition. *Language Learning, 33*(1), 55-76.

Ueno, M., & Garnsey, S. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes, 23*(5), 646-688.

Veenstra, T., & Besten, J. (1994). Fronting. In P. M. J. Arends & N. Smith (Eds.), *Pidgins and Creoles* (pp. 303-315). Amsterdam: John Benjamins Publishing Company.

Velupillai, V. (2012). *An introduction to linguistic typology*. Amsterdam: John Benjamins Publishing Company.

Versteegh, K. (2001). *The Arabic language*. Edinburgh: Edinburgh University Press.

Willis, D. (2011). The limits of resumption in Welsh Wh-dependencies. In A. Rouveret (Ed.), *Resumptive pronouns at the interfaces* (pp. 189-222). Amsterdam: John Benjamins Publishing Company.

Witzlack-Makarevich, A. (2010). *Typological variations in grammatical relations.* Unpublished doctoral dissertation University of Leipzig, Leipzig.

Wu, T. (2011). The syntax of prenominal relative clauses: A typological study. *Linguistic Typology, 15*(3), 569-623.

Yip, V., & Matthews, S. (2007). Relative clauses in Cantonese-English bilingual children: Typological challenges and processing motivations. *Studies in Second Language Acquisition, 29*(2), 277-300.

Appendix A: Types of OBL and GEN Relative Clauses

A. Types of OBL Relative Clauses:

- **Place**

1. ...wa al-mawḏiʿ allaḏī wulida fī-hi
   ...and the-place REL(3.SG.MAS) born in-it
   ...and the place in which he was born.
   [CA, 4: 81]

- **Time**

2. wa mata fī al-yawmi allaḏī mata fī-
   and died in the-day REL(3.SG.MAS) died in-
   -hi Abu Bakr
   -it Abu Bakr
   And he died in the day in which Abu Bakr died.
   [CA, 5: 305]

- **Mean**

3. ama ʿamal yatašābah bi-hi bi-al-ḥayawān...
   either act resemble in-it in-the-animal...
   Either an act by which he resembles the animal...
   [CA, 3: 83]

- **Manner**

4. ..wa šaqqa-hu ʿala al-ṣifah allatī
   ..and cut-it on the-way REL(3.SG.FEM)
   šaqqa al-ḍabiy-at
   cut the-dear-FEM
   ...and he cut it in the same way by which he cut the female dear.
   [CA, 6: 40]

- **Condition**

5. kān-at al-ǧuṯaṯ qad taḫašab-at fi
was-FEM the-dead bodies indeed harden-FEM in
waḍʿ-i-hā allaḏī dāham-hā fi-hi al-
position-GEN-its REL(3.SG.MAS) surprise-them in-it the
ḥarīq
fire
The dead bodies were hardened in the positions in which they were when the fire started.
[MSA, 2: 240]

- **Possession**

6. fatā la-hu mustaqbal mušriq
young man to-him future bright
A young man who has a bright future.
[MSA, 3: 22]

- **Source**

7. al-bāb al-yiǧī-k minn-ah rīḥ sidd-ah wa
the-door REL-come-you from-it wind close-it and
istrīḥ
rest
Close the door from which the wind comes to you, and have a rest.
[IA, 3: 106]

- **Cause**

8. al-muškil-āt allatī ʿānā min-hā al-
the-problem-PL REL(3.SG.FEM) suffered from-it the-
-muǧtamaʿ al-Muslim
-society the-Muslim
The problems from which the Muslim society suffered. [MSA, 5: 189]

- **Accompaniment**

9. halaka al-maliku allaḏī kān maʿa-hu al-rayan bin al-waleed
   died the-king REL(3.SG.MAS) was with-him Al-rayan son Al-waleed
   The king who was with Al-rayan son of Al-waleed died.
   [CA, 3: 130]

- **Concern**

10. al-šī illi tfakir bī-h gāʿid yfakrūn bī-h al-masʾūl-īn
    the-thing REL think in-it are thinking in-it the-official-PL
    The officials are thinking about the same thing which you think about.
    [IA, 3: 28]

- **Result**

11. yuaʾkid ṣiḥat al-ʾistintāǧ allaḏī tawaṣl-nā ilai-hi
    verify rightness the-conclusion REL(3.SG.MAS) came-we to-it
    Confirmed the rightness of the conclusion which we came up with.
    [MSA, 5: 86]

- **Basis**

12. yataḥadaṯ ʿan al-naqṣ allaḏī taʾsasa-t ʿalai-hi al-bašarīya
    talk about the-deficiency REL(3.SG.MAS) founded-FEM on-it the-mankind
    He was talking about the deficiency upon which the mankind was founded.
    [MSA, 1: 162]

Types of GEN Relative Clauses

- **Possession**

In this kind of GEN relative clauses, the head is the possessor.

1. taḥta-hu baġlat la-hu qad quṭiʿa
   under-him mule-FEM for-him indeed cut-PASS
   ḏanabu-*hā*
   tail-*its*
   He was on his female mule whose tail was cut.
   [CA, 4: 1107]

- **Patient**

In this kind of GEN relative clauses, the head is the patient of the action where the first part of the GEN structure is either verbal noun or active participle, as in the following example:

2. fa-ṯabat-ū fī al-manāṭiq allatī tamma
   so-steady-they in the-places REL(3.SG.MAS) done
   fatḥa-ha
   opening-their
   So they were steady in the places which were already conquered.
   [MSA, 4: 51]

- **Identity**

In this kind of GEN relative clauses, the head is identified or explained by the second part of the GEN structure, which is found in the relative clause.

3. qatla aṣhaba-hu qatlan lam tasmaʿ-ū bi-miṯli-hi
   killed friends-his killing not you.hear-PL in-similar-its
   He killed his friends in a way which you never heard of its similar (whose no similar). [CA, 4: 424]

- **Agent**

In this kind of GEN relative clauses, the head is the agent while the other part of the GEN structure inside the relative clause, is a verbal noun.

4. ġair ābih bi-ʾafruʿ al-šağar allatī
   not care in-branches the-trees REL(3.SG.FEM)
   yatawālā ṣuqūṭi-hā
   do successively falling-its
   He does not care to the trees' branches whose falling is successive (which are falling successively).
   [MSA, 2: 14]

- **Adverbial**[32]

In this kind of GEN relative clauses, the head is a place or time while the other part of GEN is prepositional[33].

5. ...ʿinda al-ṣaḫrah allatī yanfalitu ʿinda-hā
   ...at the-rock REL(3.SG.FEM) escape at-it
   al-ḥūt
   the-whale
   ...at the rock where the whale escapes.
   [CA, 4: 126]

[32] Adverbial genitive is included under descriptive genitive (see Janet and Watson (1993))

[33] A distinction is made in the literature between prepositions and prepositional in spite of the similarity between them in referring to time and place. Whereas prepositions are underived elements which are referred to as particles, prepositional are considered as nouns that are referred to in Arabic as dhruf 'qualifiers' (Badawi et al., 2004; Ryding, 1977, 2005). Apart from their adverbial function, prepositionals can retain their noun status (Badawi et al., 2004, p. 174). Therefore, they can be used in annexation structures (i.e. the GEN relative clauses). In this thesis, relative clauses with prepositionals are included under the GEN position not under the OBL position.

- **Partitive**

In this kind of GEN relative clauses, the first part of the GEN structure, which can be a quantifier or superlative, identifies a part or the quantity of the second term.

6. ...taṣrīf al-ʾala-āt allatī yuḥārbu bi-baʿḍi
   ...ordinance the-tool-PL REL(3.SG.FEM) fight by-some-
   -hā
   -them
   ...ordinance of the tools with some of which he fights...
   [CA, 2: 41]

- **Reason**

This kind of GEN relative clauses refers to the reason for which something happened.

7. wa zāla al-sabab allaḏī kun-nā
   and disappeared the-reason REL(3.SG.MAS) were-we
   nuḥbas u li-ʾağli-hi
   we.imprison.PASS for-sake-its
   The reason for which we were imprisoned ceased to exist.
   [CA, 1: 323]

Appendix B: Statistical Models Used in This Study

Table 1: Mixed Effect Poisson Regression for the Distribution of Relative Clauses in Arabic Texts in Comparison to OBL Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.719	0.188	3.819	0.000*
TSU	0.185	0.164	1.128	0.259
ISU	0.547	0.127	4.300	<0.001*
DO	0.510	0.118	4.325	<0.001*
GEN	-1.027	0.136	-7.562	<0.001*
DEF	0.987	0.043	23.105	<0.001*
INANIM	1.059	0.043	24.379	<0.001*
IA	-1.273	0.223	-5.698	<0.001*
MSA	-0.279	0.127	-2.205	0.027*
NONFIC	0.541	0.127	4.280	<0.001*

Table 2: Distribution of Relative Clauses in Arabic Texts in Comparison to TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.904	0.243	3.724	0.000*
ISU	0.362	0.185	1.953	0.051.
DO	0.325	0.104	3.117	<0.001*
OBL	-0.185	0.164	-1.128	0.259
GEN	-1.212	0.194	-6.258	<0.001*
DEF	0.987	0.043	23.105	<0.001*
INANIM	1.059	0.043	24.379	<0.001*
IA	-1.273	0.223	-5.698	<0.001*
MSA	-0.279	0.127	-2.205	0.027*
NONFIC	0.541	0.127	4.281	<0.001*

Table 3: Distribution of Relative Clauses in Arabic Texts While Combing the TSU and ISU Relative Clauses.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.674	0.145	11.563	0.000*
DO	-0.608	0.078	-7.782	<0.001*
OBL	-1.128	0.109	-10.336	<0.001*
GEN	-2.161	0.145	-14.921	<0.001*
DEF	0.986	0.043	23.093	<0.001*
INANIM	1.061	0.044	24.400	<0.001*
IA	-1.009	0.224	-4.504	<0.001*
MSA	0.292	0.149	1.958	0.050.
NONFIC	0.270	0.149	1.815	0.070.

Table 4: Distribution of anchors in TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.142	0.269	4.254	<0.001*
3rd pronoun	0.201	0.159	1.262	0.207
Def.NP	1.584	0.130	12.232	<0.001*
IA	-0.904	0.407	-2.221	0.026*
MSA	1.004	0.282	3.558	0.000*
NONFIC	-0.186	0.282	-0.661	0.508

Table 5: Distribution of Anchors in DO Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.464	0.184	13.418	<0.001*
3rd pronoun	-0.010	0.098	-0.098	0.922
Def.NP	0.399	0.089	4.471	<0.001*
IA	-0.838	0.286	-2.931	0.003*
MSA	0.449	0.199	2.259	0.024*
NONFIC	0.043	0.198	0.216	0.829

Table 6: Distribution of anchors in OBL Relative Clauses.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.400	0.324	4.323	<0.001*
3rd pronoun	0.758	0.148	5.115	<0.001*
Def.NP	1.104	0.141	7.819	<0.001*
IA	-1.623	0.514	-3.159	0.002*
MSA	0.234	0.345	0.676	0.499
NONFIC	0.038	0.345	0.111	0.912

Table 7: Distribution of Definite Relative Clauses in Comparison with the ISU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.492	0.245	6.080	<0.001*
TSU	-0.366	0.221	-1.659	0.097 .
DO	0.100	0.153	0.653	0.514
OBL	-0.511	0.131	-3.901	<0.001*
GEN	-1.971	0.229	-8.591	<0.001*
DEF	0.987	0.043	23.105	<0.001*
INANIM	1.199	0.053	22.728	<0.001*
IA	-0.492	0.349	-1.408	0.159
MSA	0.644	0.210	3.061	0.002 *
NONFIC	0.487	0.207	2.359	0.018*

Table 8: Distribution of Indefinite Relative Clauses in Comparison with the DO Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.355	0.342	3.958	<0.001*
TSU	0.145	0.145	0.998	0.319
ISU	0.267	0.223	1.194	0.233
OBL	-0.278	0.170	-1.634	0.102
GEN	-0.578	0.238	-2.426	0.015 *
DEF	0.987	0.043	23.105	<0.001*
INANIM	0.723	0.078	9.299	<0.001*
IA	-1.413	0.434	-3.256	0.001 *
MSA	-0.642	0.273	-2.346	0.019 *
NONFIC	0.075	0.264	0.282	0.778

Table 9: Distribution of Definite Relative Clauses in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.982	0.298	3.295	0.001*
TSU	0.145	0.236	0.613	0.540
ISU	0.511	0.131	3.901	<0.001*
DO	0.610	0.171	3.562	0.000*
GEN	-1.460	0.190	-7.686	<0.001*
DEF	0.987	0.043	23.105	<0.001*
INANIM	1.199	0.053	22.729	<0.001*
IA	-0.492	0.349	-1.408	0.159
MSA	0.644	0.210	3.061	0.002*
NONFIC	0.487	0.207	2.359	0.018*

Table 10: Distribution of Indefinite Relative Clauses in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.077	0.298	3.616	0.000*
TSU	0.421	0.122	3.446	0.001*
ISU	0.538	0.230	2.341	0.019*
DO	0.269	0.160	1.681	0.093 .
GEN	-0.306	0.227	-1.347	0.178
DEF	0.987	0.043	23.105	<0.001*
INANIM	0.723	0.078	9.300	<0.001*
IA	-1.403	0.432	-3.245	0.001*
MSA	-0.629	0.273	-2.304	0.021*
NONFIC	0.0713	0.263	0.271	0.787

Table 11: Difference between Definite and Indefinite Relative Clauses in the TSU position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.016	0.254	3.994	<0.001*
DEF	0.939	0.091	10.336	<0.001*
INANIM	0.571	0.085	6.715	<0.001*
IA	-0.893	0.393	-2.274	0.023*
MSA	0.958	0.276	3.468	0.001*
NONFIC	-0.208	0.276	-0.753	0.451

Table 12: Difference between Definite and Indefinite Relative Clauses in the ISU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.713	0.162	10.609	<0.001*
DEF	0.882	0.078	11.263	<0.001*
INANIM	0.451	0.073	6.172	<0.001*
IA	-1.392	0.264	-5.264	0.023*
MSA	-0.082	0.166	-0.497	0.620
NONFIC	0.464	0.166	2.798	0.005*

Table 13: Difference between Definite and Indefinite Relative Clauses in the DO Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.291	0.206	1.410	0.15842
DEF	1.233	0.088	14.096	<0.001*
INANIM	1.803	0.105	17.183	<0.001*
IA	-0.826	0.283	-2.916	0.004 *
MSA	0.475	0.197	2.409	0.016*
NONFIC	0.014	0.197	0.071	0.944

Table 14: Difference between Definite and Indefinite Relative Clauses in the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.691	0.308	-2.247	0.025*
DEF	1.188	0.109	10.941	<0.001*
INANIM	2.733	0.192	14.245	<0.001*
IA	-1.593	0.405	-3.931	<0.001*
MSA	0.039	0.271	0.142	0.887
NONFIC	-0.087	0.271	-0.320	0.749

Table 15: Difference between Definite and Indefinite Relative Clauses in the GEN Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.610	0.268	2.277	0.023 *
DEF	0.164	0.160	1.024	0.306
INANIM	0.637	0.168	3.800	<0.001*
IA	-1.041	0.416	-2.503	0.012*
MSA	-0.034	0.265	-0.128	0.898
NONFIC	0.006	0.265	0.021	0.983

Table 16: Distribution of the Grammatical Positions, in Comparing with OBL, that are Occupied by Heads of ISU Relative Clauses in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.439	0.166	8.680	<0.001*
SU	-0.789	0.125	-6.341	<0.001*
TOPIC	-1.206	0.145	-8.319	<0.001*
DO	-0.264	0.106	-2.499	0.013*
GEN	-0.391	0.110	-3.570	<0.001 *
OTHERS	-2.768	0.286	-9.670	<0.001*
SUCOM	-1.526	0.165	-9.269	<0.001*
DEF	0.861	0.082	10.508	<0.001*
IA	-1.405	0.261	-5.392	<0.001*
MSA	-0.108	0.164	-0.656	0.512
NONFIC	0.439	0.164	2.677	0.007*

Table 17: Distribution of the Grammatical Positions, in Comparing with OBL, that are Occupied by Heads of TSU Relative Clauses in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.843	0.276	3.056	0.002*
SU	-1.065	0.163	-6.556	<0.001*
TOPIC	-0.920	0.154	-5.968	<0.001*
DO	-0.235	0.124	-1.898	0.058.
GEN	-0.334	0.127	-2.621	0.009*
OTHERS	-2.164	0.256	-8.443	<0.001*
SUCOM	-1.147	0.168	-6.845	<0.001*
DEF	0.859	0.094	9.128	<0.001*
IA	-0.958	0.418	-2.295	0.022*
MSA	0.849	0.299	2.840	0.005*
NONFIC	-0.344	0.299	-1.150	0.250

Table 18: Distribution of the Grammatical Positions, in Comparing with OBL, that are Occupied by DO Relative Clauses in Arabic texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.051	0.214	4.903	<0.001*
SU	-1.458	0.168	-8.701	<0.001*
TOPIC	-1.022	0.142	-7.223	<0.001*
DO	-0.231	0.109	-2.112	0.035*
GEN	-0.225	0.109	-2.055	0.040*
OTHERS	-3.045	0.342	-6.767	<0.001*
SUCOM	-0.924	0.137	-6.845	<0.001*
DEF	1.175	0.090	13.080	<0.001*
IA	-0.880	0.310	-2.839	0.005*
MSA	0.377	0.222	1.697	0.090.
NONFIC	-0.100	0.223	-0.448	0.654

Table 19: Distribution of the Grammatical Positions, in Comparison to OBL, that are Occupied by OBL Relative Clauses in Arabic texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.279	0.266	4.815	<0.001*
SU	-1.978	0.210	-9.443	<0.001*
TOPIC	-1.710	0.187	-9.165	<0.001*
DO	-1.002	0.141	-7.112	0.035*
GEN	-0.842	0.133	-6.327	0.040*
OTHERS	-2.752	0.298	-9.230	<0.001*
SUCOM	-1.626	0.180	-9.027	<0.001*
DEF	1.164	0.111	10.482	<0.001*
IA	-1.618	0.416	-3.889	<0.001*
MSA	-0.011	0.284	-0.037	0.971
NONFIC	-0.138	0.284	-0.484	0.628

Table 20: Distribution of the Grammatical Positions, in Comparison with OBL, that are Occupied by GEN Relative Clauses in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.265	0.290	0.912	0.362
SU	-0.419	0.259	-1.617	0.106
TOPIC	-1.692	0.414	-4.090	<0.001*
DO	-0.204	0.243	-0.837	0.403
GEN	-0.082	0.236	-0.349	0.727
OTHERS	-1.558	0.391	-3.983	<0.001*
SUCOM	-1.073	0.323	-3.320	<0.001*
DEF	0.146	0.162	0.905	0.365
IA	-1.049	0.425	-2.467	0.014 *
MSA	-0.042	0.275	-0.152	0.879
NONFIC	0.034	0.275	0.125	0.900

Table 21: Distribution of Relative Clauses with Heads Occupy the OBL Position in Arabic Texts.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.301	0.148	8.787	<0.001*
TSU	-0.584	0.227	-2.575	0.010 *
DO	-0.142	0.123	-1.160	0.246
OBL	-0.248	0.160	-1.553	0.120
GEN	-1.806	0.211	-8.566	<0.001*
DEF	1.270	0.087	14.640	<0.001*
IA	-2.000	0.309	-6.478	0.001 *
MSA	-0.053	0.148	-0.361	0.718
NONFIC	0.153	0.147	1.036	0.300

Table 22: Distribution of Relative Clauses with Heads Occupy the 'GEN' Position in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.259	0.262	-0.987	0.324
TSU	-0.188	0.216	-0.867	0.386
DO	0.296	0.254	1.168	0.243
OBL	-0.475	0.212	-2.239	0.025*
GEN	-1.705	0.301	-5.668	<0.001*
DEF	1.517	0.113	13.447	<0.001*
IA	-1.025	0.310	-3.303	0.001*
MSA	1.241	0.144	8.614	0.718
NONFIC	0.164	0.128	1.280	0.201

Table 23: Distribution of Relative Clauses with Heads Occupy the 'DO' Position in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.577	0.222	7.093	<0.001
TSU	-0.171	0.253	-0.676	0.499
DO	-0.013	0.237	-0.053	0.958
OBL	-0.899	0.228	-3.951	<0.001*
GEN	-1.602	0.326	-4.914	<0.001*
DEF	0.320	0.089	3.597	<0.001*
IA	-1.734	0.372	-4.664	<0.001*
MSA	-0.275	0.220	-1.248	0.212
NONFIC	0.145	0.218	0.663	0.507

Table 24: Distribution of Relative Clauses with Heads Occupy the 'SU' Position in Arabic Texts.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.104	0.301	3.663	<0.001
TSU	-0.630	0.299	-2.103	0.035*
DO	-0.973	0.422	-2.305	0.958
OBL	-1.23	0.228	3.877	<0.001*
GEN	-0.314	-4.073	-4.914	<0.001*
DEF	-0.083	0.133	-0.621	0.535
IA	-1.703	0.522	-3.266	<0.001*
MSA	-1.025	0.288	-3.557	<0.001*
NONFIC	1.078	0.281	3.832	<0.001*

Table 25: Distribution of Relative Clauses with Heads Occupy the 'SU(TOPIC)' Position in Arabic Texts.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.249	0.469	-4.796	<0.001
TSU	-0.383	0.273	-1.405	0.160
DO	0.168	0.246	0.683	0.494
OBL	-0.515	0.229	-2.250	0.024*
GEN	-3.138	0.707	-4.439	<0.001*
DEF	3.411	0.366	9.331	<0.001*
IA	-0.224	0.412	-0.544	0.586
MSA	0.480	0.291	1.649	0.099.
NONFIC	-0.200	0.287	-0.695	0.487

Table 26: Distribution of Relative Clauses with Heads Occupy the 'SUCOM' Position in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.327	0.445	-0.734	0.463
TSU	0.065	0.269	0.240	0.811
DO	0.546	0.239	2.282	0.023*
OBL	-0.321	0.277	-1.158	0.247
GEN	-1.416	0.442	-3.203	0.001*
DEF	0.947	0.152	6.237	<0.001*
IA	-0.118	0.598	-0.197	0.844
MSA	0.254	0.462	0.549	0.583
NONFIC	-0.587	0.465	-1.263	0.207

Table 27: Distribution of Relative Clauses with Heads Occupy the 'OTHERS' Positions in Arabic Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.346	0.539	-0.643	0.520
TSU	0.681	0.453	1.503	0.133
DO	-0.174	0.670	-0.260	0.795
OBL	0.144	0.551	0.261	0.794
GEN	-0.275	0.621	-0.442	0.658
DEF	-1.338	0.334	-4.002	<0.001*
IA	-0.939	0.545	-1.722	0.085.
MSA	-0.949	0.438	-2.166	0.030*
NONFIC	-0.313	0.410	-0.764	0.445

Table 28: Distribution of Relative Clauses with Heads Occupy Different Grammatical Positions in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.054	0.135	7.779	< 0.001 *
TSU	-0.411	0.196	-2.101	0.036*
DO	-0.080	0.148	-0.543	0.587
OBL	-0.560	0.135	-4.156	< 0.001 *
GEN	-0.025	0.062	-0.403	0.687
GEN HEADS	-0.810	0.120	-6.769	< 0.001 *
OBL HEADS	0.381	0.057	6.735	0.015 *
OTHERS HEADS	-2.188	0.137	-10.971	< 0.001 *
SUCOM HEADS	-0.885	0.081	-9.054	< 0.001 *
SU HEADS	0.078	0.110	-10.070	< 0.001 *
TOPIC HEADS	-0.827	0.080	-10.462	< 0.001 *
DEF	0.932	0.044	21.107	0.000*
NONFIC	0.515	0.132	3.915	< 0.001 *
IA	-1.250	0.223	-5.611	< 0.001 *
MSA	-0.229	0.132	-1.740	0.082.

Table 29: Distribution of Relative Clauses Referring to Animate Referents in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.838	0.288	6.393	<0.001*
TSU	-0.166	0.205	-0.807	0.420
DO	-0.846	0.216	-3.924	<0.001*
OBL	-2.091	0.266	-7.858	<0.001*
GEN	-1.462	0.303	-4.831	<0.001*
DEF	0.643	0.079	8.148	<0.001*
IA	-0.768	0.333	-2.308	0.021 *
MSA	-0.311	0.237	-1.312	0.189
NONFIC	0.031	0.237	0.133	0.894

Table 30: Distribution of Relative Clauses Referring to Animate Referents in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.664	0.256	6.508	<0.001*
ISU	0.171	0.207	0.826	0.409
DO	-0.681	0.142	-4.799	<0.001*
OBL	-1.911	0.262	-7.285	<0.001*
GEN	-1.298	0.244	-5.326	<0.001*
DEF	0.643	0.079	8.148	<0.001*
IA	-0.786	0.332	-2.367	0.018*
MSA	-0.328	0.237	-1.384	0.167
NONFIC	0.078	0.236	0.330	0.741

Table 31: Distribution of Relative Clauses Referring to Animate Referents in Comparison with GEN

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.367	0.255	1.440	0.150
ISU	1.468	0.302	4.861	<0.001
TSU	1.297	0.244	5.321	<0.001*
DO	0.617	0.233	2.643	0.008*
OBL	-0.614	0.271	-2.265	0.024*
DEF	0.643	0.079	8.148	<0.001*
IA	-0.786	0.332	-2.369	0.018*
MSA	-0.328	0.237	-1.386	0.166
NONFIC	0.077	0.236	0.326	0.745

Table 32: Distribution of Relative Clauses Referring to Inanimate Referents in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.651	0.132	20.083	<0.001
ISU	1.468	0.302	4.861	<0.001
TSU	1.297	0.281	-2.271	0.023*
DO	0.304	0141	2.152	0.031*
OBL	-0.140	0.146	-0.961	0.337
GEN	-1.566	0.168	-9.307	<0.001*
DEF	0.643	0.079	8.148	<0.001*
IA	-1.177	0.227	-5.180	<0.001*
MSA	-0.015	0.118	-0.123	0.902
NONFIC	0.405	0.116	3.502	<0.001*

Table 33: Distribution of Relative Clauses Referring to Inanimate Referents in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.305	0.337	3.875	<0.001*
ISU	0.638	0.281	2.271	0.023*
DO	0.942	0.243	3.876	<0.001*
OBL	0.498	0.296	1.680	0.093.
GEN	-0.928	0.265	-3.509	<0.001*
DEF	1.119	0.051	21.868	<0.001*
IA	-1.178	0.228	-5.178	<0.001*
MSA	-0.016	0.118	-0.137	0.891
NONFIC	0.407	0.116	3.518	<0.001*

Table 34: Interaction between Definiteness and Animacy of the Heads of the Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.504	0.139	10.783	<0.001*
TSU	-0.362	0.185	-1.953	0.051 .
DO	-0.036	0.131	-0.279	0.780
OBL	-0.547	0.127	-4.300	<0.001*
GEN	-1.573	0.164	-9.577	<0.001*
DEF	1.119	0.051	21.868	<0.001*
IA	-1.273	0.223	-5.699	<0.001*
MSA	-0.279	0.127	-2.205	0.027*
NONFIC	0.542	0.127	4.281	<0.001*
INANIM	0.723	0.078	9.305	<0.001*
DEF	0.643	0.079	8.164	<0.001*
INANIM:DEF	0.476	0.094	5.064	<0.001*

Table 35: Interaction between variety (MSA and IA) and relative clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.801	0.249	-3221	0.001 *
TSU	0.012	0.296	0.040	0.968
DO	0.444	0290	1.528	0.126
OBL	-0.408	0.378	-1.270	0.204
GEN	-1.097	0.508	-2.158	0.030 *
MSA	1.554	0.272	5.704	< 0.001*
DEF	1.573	0.066	23.843	< 0.001*
INANIM	1.284	0.060	21.261	< 0.001*
TSU:MSA	0.133	0.331	0.401	0.688
DO:MSA	-0.260	0.330	-0.787	0.431
OBL:MSA	0.006	0.432	0.015	0.988
GEN:MSA	-0.546	0.593	-0.921	0.357

Table 36: Distribution of Relative Clauses in CA texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.582	0.228	6.934	<0.001*
TSU	-0.986	0.278	-3.543	<0.001*
DO	-0.375	0.176	-2.129	0.033*
OBL	-0.628	0.184	-3.416	<0.001*
GEN	-1.680	0.183	-9.198	<0.001*
DEF	0.359	0.060	6.000	<0.001*
INANIM	0.780	0.063	12.310	<0.001*
NONFIC	0.964	0.202	4.768	<0.001*

Table 37: Frequency of Relative Clauses in Comparison with OBL Relative Clauses in CA Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.943	0.378	2.496	0.013*
TSU	-0.363	0.219	-1.658	0.097.
DO	0.255	0.217	1.173	0.241
ISU	0.625	0.183	3.410	<0.001*
GEN	-1.059	0.234	-4.522	<0.001*
DEF	0.359	0.060	6.000	<0.001*
INANIM	0.780	0.063	12.310	<0.001*
NONFIC	0.989	0.203	4.865	<0.001*

Table 38: Distribution of Relative Clauses Including Absolutive as a Category in CA Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.171	0.214	10.136	<0.001*
ABS	0.185	0.058	3.180	0.002*
DO	-0.749	0.113	-6.646	<0.001*
OBL	-0.995	0.171	-5.834	<0.001*
GEN	-2.032	0.134	-15.175	<0.001*
DEF	0.390	0.048	8.182	<0.001*
INANIM	0.744	0.050	14.860	<0.001*
NONFIC	0.512	0.220	2.325	0.020*

Table 39: Distribution of Definite Relative Clauses in CA Texts in Comparison to ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.028	0.417	4.855	<0.001*
TSU	-1.055	0.342	-3.085	0.002*
DO	-0.214	0.275	-0.779	0.436
OBL	-0.582	0.203	-2.873	0.004*
GEN	-2.319	0.409	-5.671	<0.001*
INANIM	0.904	0.085	10.674	<0.001*
NONFIC	0.269	0.285	0.944	0.345

Table 40: Distribution of Definite Relative Clauses in CA Texts in Comparison with DO

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.814	0.299	6.073	<0.001*
TSU	-0.841	0.197	-4.267	<0.001*
ISU	0.214	0.275	-0.779	0.436
OBL	-0.367	0.364	-1.009	0.313
GEN	-2.104	0.317	-6.645	<0.001*
INANIM	0.904	0.085	10.674	<0.001*
NONFIC	0.269	0.285	0.944	0.345

Table 41: Distribution of Definite Relative Clauses in CA Texts in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.973	0.357	2.725	0.006*
DO	0.841	0.197	4.267	<0.001*
ISU	1.055	0.342	3.085	0.002*
OBL	0.473	0.405	1.170	0.242
GEN	-1.264	0.255	-4.947	<0.001*
INANIM	0.904	0.085	10.674	<0.001*
NONFIC	0.269	0.285	0.944	0.345

Table 42: Distribution of Indefinite Relative Clauses in CA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.315	0.453	2.903	0.004*
DO	-0.861	0.582	-1.478	0.139
TSU	-0.914	0.590	-1.550	0.121
OBL	-0.933	0.458	-2.037	0.042*
GEN	-1.018	0.218	-4.662	<0.001*
INANIM	0.612	0.096	6.368	<0.001*
NONFIC	1.070	0.291	3.683	<0.001

Table 43: Distribution of Indefinite Relative Clauses in CA texts in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.401	0.952	0.421	0.673
DO	0.054	0.196	0.274	0.784
ISU	0.914	0.590	1.550	0.121
OBL	-0.018	0.263	-0.070	0.944
GEN	-0.104	0.674	-0.154	0.877
INANIM	0.612	0.096	6.368	<0.001*
NONFIC	1.070	0.291	3.683	<0.001*

Table 44: Distribution of Indefinite Relative Clauses in CA Texts in Comparison with OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.365	0.799	0.457	0.648
DO	0.131	0.278	0.473	0.636
ISU	0.979	0.468	2.092	0.036*
TSU	0.183	0.170	1.076	0.282
GEN	-0.022	0.539	-0.041	0.968
INANIM	0.612	0.096	6.368	<0.001*
NONFIC	1.000	0.306	3.270	0.001*

Table 45: Distribution of Relative Clauses Referring to Animate Referents in CA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.936	0.378	5.126	<0.001*
DO	-1.139	0.356	-3.200	0.001*
TSU	-0.593	0.296	-2.003	0.045*
OBL	-2.089	0.465	-4.490	<0.001*
GEN	-1.463	0.422	-3.470	<0.001*
DEF	0.160	0.105	1.519	0.129
NONFIC	0.599	0.353	1.699	0.089.

Table 46: Distribution of Relative Clauses Referring to Animate Referents in CA in Comparison to DO

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.798	0.371	2.148	0.032*
ISU	1.139	0.356	3.200	0.001*
TSU	0.545	0.182	2.991	0.003*
OBL	-0.950	0.352	-2.700	0.007*
GEN	-0.324	0.358	-0.906	0.365
DEF	0.160	0.105	1.519	0.129
NONFIC	0.599	0.353	1.699	0.089.

Table 47: Distribution of Relative Clauses Referring to Animate Referents in CA Texts in Comparison with GEN

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.388	0.325	1.195	0.232
ISU	1.478	0.421	3.511	0.001.*
TSU	0.873	0.344	2.537	0.011*
DO	0.328	0.374	0.877	0.380
OBL	-0.636	0.328	-1.937	0.0523.
DEF	0.160	0.105	1.519	0.129
NONFIC	0.753	0.3567	2.110	0.0345 *

Table 48: Distribution of Relative Clauses Referring to Inanimate Referents in CA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.338	0.245	9.531	<0.001*
TSU	-1.444	0.482	-2.996	0.003*
DO	0.0317	0.202	0.157	0.875
OBL	-0.173	0.2144	-0.806	0.421
GEN	-1.642	0.220	-7.460	<0.001*
DEF	0.452	0.073	6.203	<0.001*
NONFIC	0.463	0.219	2.115	0.034*

Table 49: Distribution of Relative Clauses Referring to Inanimate Referents in CA Texts in Comparison with OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.166	0.367	5.908	<0.001*
TSU	-1.271	0.542	-2.346	0.019*
ISU	0.173	0.214	0.806	0.421
DO	0.204	0.284	0.721	0.471
GEN	-1.470	0.273	-5.392	<0.001*
DEF	0.452	0.073	6.203	<0.001*
NONFIC	0.463	0.219	2.115	0.034*

Table 50: Distribution of Relative Clauses Referring to Inanimate Referents in CA Texts in Comparison to GEN

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.696	0.311	2.237	0.025*
TSU	0.199	0.451	0.440	0.660
ISU	1.642	0.220	7.460	<0.001*
DO	1.674	0.179	9.345	<0.001*
OBL	1.470	0.273	5.392	<0.001*
DEF	0.452	0.073	6.203	<0.001*
NONFIC	0.463	0.219	2.115	0.034*

Table 51: Distribution of Relative Clauses in Fiction CA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.837	0.321	2.605	0.009*
TSU	-0.260	0.170	-1.528	0.126
DO	0.098	0.188	0.519	0.604
OBL	-0.493	0.164	-2.999	0.003*
GEN	-1.261	0.214	-5.885	<0.001*
DEF	0.517	0.077	6.203	<0.001*
INANIM	1.121	0.087	6.708	<0.001*

Table 52: Distribution of Relative Clauses in CA Fiction Texts in Comparison with OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.344	0.440	0.782	0.434
TSU	0.233	0.2416	0.963	0.335
ISU	0.493	0.1644	2.999	0.003*
DO	0.591	0.289	2.045	0.041*
GEN	-0.768	0.296	-2.590	0.010*
DEF	0.517	0.077	6.707	<0.001*
INANIM	1.121	0.087	12.951	<0.001*

Table 53: Distribution of Relative Clauses in CA Nonfiction Texts in Comparison with DO

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.990	0.400	4.975	<0.001*
TSU	-0.932	0.180	-5.194	<0.001*
ISU	0.583	0.240	2.434	0.015*
OBL	-0.289	0.289	-1.863	0.062.
GEN	-1.294	0.198	-6.521	<0.001*
DEF	0.297	0.082	3.611	<0.001*
INANIM	0.331	0.083	4.014	<0.001*

Table 54: Distribution of Relative Clauses in CA Nonfiction Texts in Comparison to TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.058	0.512	2.065	0.039*
DO	0.932	0.180	5.194	<0.001*
ISU	1.515	0.351	4.314	<0.001*
OBL	0.643	0.251	2.560	0.011*
GEN	-0.362	0.278	-1.301	0.193
DEF	0.297	0.082	3.611	<0.001*
INANIM	0.331	0.083	4.014	<0.001*

Table 55: Differences among CA Relative Clauses According to the Positions of their Heads in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.547	0.234	6.618	<0.001*
TSU	-0.988	0.278	-3.553	<0.001*
DO	-0.371	0.177	-2.102	0.036*
OBL	-0.624	0.183	-3.407	0.001*
GEN	-1.661	0.183	-9.094	<0.001*
DO HEADS	-0.224	0.079	-2.827	0.005*
GEN HEADS	-0.853	0.097	-8.835	<0.001*
OTHERS HEADS	-2.328	0.177	-13.148	<0.001*
SUCOM HEADS	-1.395	0.118	-11.779	<0.001*
SU HEADS	-0.859	0.097	-8.882	<0.001*
TOPIC HEADS	-1.477	0.122	-12.065	<0.001*
NONFIC	1.008	0.205	4.921	<0.001*
DEF	0.355	0.060	5.937	<0.001*

Table 56: Distribution of Relative Clauses in MSA Texts Compared to ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.493	0.156	3.102	0.002*
DO	0.186	0.164	1.135	0.256
TSU	0.127	0.118	1.080	0.280
OBL	-0.471	0.219	-2.155	0.031*
GEN	-1.636	0.318	-5.142	<0.001*
DEF	1.684	0.072	23.274	<0.001*
INANIM	1.408	0.066	21.279	<0.001*
NONFIC	0.129	0.055	2.364	0.018*

Table 57: Distribution of Relative Clauses in MSA Texts in Comparison to the OBL position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.068	0.110	0.623	0.533
ISU	0.457	0.154	2.972	0.003*
TSU	0.584	0.099	5.902	<0.001*
DO	0.643	0.086	7.514	<0.001*
GEN	-1.178	0.219	-5.378	<0.001*
DEF	1.684	0.072	23.274	<0.001*
INANIM	1.408	0.066	21.279	<0.001*
NONFIC	0.067	0.058	1.153	0.249

Table 58: Distribution of Relative Clauses Including ABS in MSA Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.230	0.111	11.067	<0.001*
ABS	0.043	0.066	0.653	0.514
DO	-0.579	0.120	-4.817	<0.001*
OBL	-1.235	0.177	-6.967	<0.001*
GEN	-2.408	0.289	-8.332	<0.001*
DEF	1.714	0.059	29.017	<0.001*
INANIM	1.392	0.053	26.167	<0.001*
NONFIC	0.158	0.051	3.119	0.002*

Table 59: Distribution of Definite Relative Clauses in MSA Texts in Comparison to the TSU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.149	0.111	19.287	<0.001*
ISU	-0.116	0.102	-1.144	0.253
DO	0.105	0.101	1.041	0.298
OBL	-0.586	0.174	-3.371	0.001*
GEN	-1.979	0.316	-6.255	<0.001*
INANIM	1.490	0.074	20.158	<0.001*
NONFIC	0.282	0.078	3.604	<0.001*

Table 60: Distribution of Definite Relative Clauses in MSA Texts in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.563	0.156	10.024	<0.001*
ISU	0.450	0.211	2.228	0.026*
DO	0.690	0.128	5.396	<0.001*
TSU	0.586	0.174	3.371	0.001*
GEN	-1.394	0.222	-6.287	<0.001*
INANIM	1.490	0.073	20.158	<0.001*
NONFIC	0.282	0.078	3.604	<0.001*

Table 61: Distribution of Indefinite Relative Clauses in MSA Texts in Comparison to the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.277	0.282	0.983	0.325
ISU	0.566	0.365	1.551	0.121
DO	0.410	0.262	1.562	0.118
TSU	0.717	0.2196	3.266	0.001*
GEN	-0.360	0.356	-1.010	0.312
INANIM	1.024	0.151	6.775	<0.001*
NONFIC	-0.068	0.218	-0.313	0.754

Table 62: Distribution of Indefinite Relative Clauses in MSA Texts in Comparison to the ISU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.845	0.348	2.430	0.015*
TSU	0.151	0.344	0.438	0.662
DO	-0.157	0.446	-0.351	0.725
OBL	-0.568	0.356	-1.596	0.111
GEN	-0.926	0.459	-2.017	0.044*
INANIM	1.024	0.151	6.775	<0.001*
NONFIC	-0.071	0.218	-0.326	0.745

Table 63: Distribution of Relative Clauses Referring to Animate Referents in MSA Texts in Comparison with the TSU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.431	0.219	6.539	<0.001*
ISU	0.032	0.282	0.114	0.910
DO	-0.858	0.264	-3.255	0.001*
OBL	-2.217	0.414	-5.352	<0.001*
GEN	-1.993	0.373	-5.335	<0.001*
DEF	1.322	0.146	9.066	<0.001*
NONFIC	-0.490	0.183	-2.677	0.007*

Table 64: Distribution of Relative Clauses Referring to Animate Referents in MSA Texts in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.758	0.331	-2.293	0.022*
ISU	2.193	0.419	5.230	<0.001*
DO	1.311	0.343	3.828	<0.001*
TSU	2.172	0.343	6.327	<0.001*
GEN	0.175	0.443	0.396	0.692
DEF	1.322	0.146	9.071	<0.001*
NONFIC	-0.435	0.182	-2.390	0.017*

Table 65: Distribution of Relative Clauses Referring to Inanimate Referents in MSA Texts in Comparison with the TSU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.783	0.146	12.253	<0.001*
ISU	-0.185	0.106	-1.749	0.080.
DO	0.288	0.163	1.763	0.078.
OBL	-0.318	0.218	-1.458	0.145
GEN	-1.646	0.262	-6.284	<0.001*
DEF	1.788	0.084	21.353	<0.001*
NONFIC	0.207	0.066	3.123	0.002*

Table 66: Distribution of Relative Clauses Referring to Inanimate Referents in MSA Texts in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.465	0.144	10.182	<0.001*
ISU	0.133	0.231	0.577	0.564
DO	0.606	0.103	5.866	<0.001*
TSU	-0.318	0.218	-1.458	0.145
GEN	-1.328	0.197	-6.736	<0.001*
DEF	1.788	0.084	21.353	<0.001*

Table 67: Distribution of Relative Clauses in MSA Fiction Texts in Comparison to the DO Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.975	0.132	7.403	<0.001*
ISU	-0.490	0.167	-2.943	0.003*
TSU	-0.147	0.127	-1.159	0.246
OBL	-0.667	0.193	-3.449	0.001*
GEN	-1.813	0.265	-6.832	<0.001*
DEF	1.604	0.104	15.475	<0.001*
INANIM	1.224	0.092	13.255	<0.001*

Table 68: Distribution of Relative Clauses in MSA Fiction Texts in Comparison to the TSU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.836	0.134	6.227	<0.001*
ISU	-0.351	0.167	-2.102	0.036*
DO	0.137	0.106	1.291	0.197
OBL	-0.528	0.193	-2.734	0.006*
GEN	-1.673	0.264	-6.345	<0.001*
DEF	1.604	0.104	15.476	<0.001*
INANIM	1.224	0.09235	13.255	<0.001*

Table 69: Distribution of Relative Clauses in MSA Fiction Texts in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.307	0.212	1.448	0.148
ISU	0.177	0.305	0.581	0.561
TSU	0.520	0.259	2.007	0.256*
DO	0.667	0.180	3.702	0.001*
GEN	-1.145	0.196	-5.834	<0.001*
DEF	1.604	0.104	15.475	<0.001*
INANIM	1.224	0.092	13.255	<0.001*

Table 70: Distribution of Relative Clauses in MSA Nonfiction Texts in Comparison with the TSU Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.558	0.136	4.106	<0.001*
ISU	0.068	0.159	0.427	0.669
DO	-0.034	0.099	-0.345	0.730
OBL	-0.688	0.120	-5.755	<0.001*
GEN	-1.905	0.363	-5.252	<0.001*
DEF	1.756	0.101	17.378	<0.001*
INANIM	1.584	0.095	16.620	<0.001*

Table 71: Distribution of Relative Clauses in MSA Nonfiction Texts in Comparison with the OBL Position

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.130	0.152	-0.857	0.392
ISU	0.757	0.173	4.370	<0.001*
TSU	0.679	0.143	4.732	<0.001*
DO	0.654	0.121	5.400	<0.001*
GEN	-1.216	0.365	-3.329	0.001*
DEF	1.756	0.101	17.378	<0.001*
INANIM	1.584	0.095	16.620	<0.001*

Table 72: Difference Among Relative Clauses According to their Positions in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.993	0.201	4.952	<0.001*
TSU	0.120	0.140	0.858	0.391
DO	0.092	0.239	0.384	0.701
OBL	-0.576	0.299	-1.925	0.054.
GEN	-1.648	0.361	-4.561	<0.001*
SU HEADS	-1.615	0.125	-12.927	<0.001*
TOPIC HEADS	-1.179	0.105	-11.240	<0.001*
DO HEADS	-0.611	0.086	-7.126	<0.001*
SUCOM HEADS	-1.528	0.121	-12.679	<0.001*
GEN HEADS	-0.127	0.074	-1.702	0.089.
OTHERS HEADS	-3.474	0.293	-11.839	<0.001*
NONFIC	0.121	0.059	2.038	0.042*
DEF	1.719	0.076	22.751	<0.001*

Table 73: Distribution of Relative Clauses in IA Texts in comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.386	0.223	1.727	0.084.
TSU	-0.133	0.400	-0.332	0.740
DO	0.397	0.260	1.523	0.128
OBL	-0.546	0.341	-1.599	0.110
GEN	-1.089	0.363	-2.998	0.003*
DEF	0.830	0.168	4.947	<0.001*
INANIM	0.435	0.158	2.755	0.006*

Table 74: Frequency of Relative Clauses in IA Texts in Comparison with OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.160	0.336	-0.478	0.633
TSU	0.414	0.287	1.444	0.149
DO	0.943	0.255	3.698	<0.001*
ISU	0.543	0.372	1.459	0.145
GEN	-0.544	0.350	-1.556	0.120
DEF	0.830	0.168	4.946	<0.001*
INANIM	0.435	0.158	2.754	0.006*

Table 75: Frequency of Relative Clauses in IA Texts while Including Absolutive as a Category in Comparison with SU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.938	0.220	4.267	<0.001*
ABS	0.246	0.167	1.472	0.141
DO	-0.279	0.176	-1.586	0.113
OBL	-1.220	0.250	-4.887	<0.001*
GEN	-1.766	0.302	-5.846	<0.001*
DEF	0.895	0.136	6.572	<0.001*
INANIM	0.564	0.129	4.389	<0.001*

Table 76: Frequency of Definite Relative Clauses in IA Texts in Comparison with OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.451	0.548	0.823	0.410
TSU	0.384	0.559	0.686	0.493
DO	1.264	0.418	3.026	0.003*
ISU	0.478	0.345	1.384	0.166
GEN	-0.987	0.559	-1.768	0.077.
INANIM	0.398	0.189	2.112	0.035*

Table 77: Frequency of Definite Relative Clauses in IA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.929	0.474	1.959	0.050.
TSU	-0.095	0.575	-0.165	0.869
DO	0.786	0.384	2.045	0.041*
OBL	-0.478	0.345	-1.384	0.166
GEN	-1.466	0.598	-2.451	0.014*
INANIM	0.398	0.189	2.112	0.035*

Table 78: Frequency of Indefinite Relative Clauses in IA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.459	1.111	-0.413	0.679
TSU	0.921	1.033	0.891	0.373
DO	0.616	0.828	0.744	0.457
OBL	0.184	0.871	0.212	0.833
GEN	-0.130	1.395	-0.093	0.926
INANIM	0.521	0.291	1.793	0.073.

Table 79: Frequency of Indefinite Relative Clauses in IA Texts in Comparison with DO

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.153	0.494	0.310	0.757
TSU	0.312	0.479	0.652	0.515
ISU	-0.659	0.878	-0.750	0.453
OBL	-0.466	0.616	-0.757	0.449
GEN	-0.742	0.802	-0.926	0.355
INANIM	0.521	0.291	1.792	0.073.

Table 80: Frequency of Relative Clauses Referring to Animate Referents in IA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.429	0.317	1.351	0.177
TSU	0.368	0.429	0.857	0.391
DO	-0.209	0.375	-0.558	0.577
OBL	-3.404	1.393	-2.443	0.015*
GEN	-0.550	0.473	-1.163	0.245
DEF	0.9057	0.274	3.301	0.001*

Table 81: Frequency of Relative Clauses Referring to Animate Referents in IA Texts in Comparison with GEN

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.041	0.371	-0.111	0.912
TSU	0.829	0.475	1.745	0.081.
ISU	0.428	0.436	0.981	0.326
DO	0.261	0.422	0.619	0.536
OBL	-2.596	1.213	3.318	0.001*
DEF	0.906	0.273	3.301	0.001*

Table 82: Frequency of Relative Clauses Referring to Inanimate Referents in IA Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.787	0.263	2.994	0.003*
TSU	0.674	0.338	1.996	0.046*
DO	0.428	0.436	0.981	0.326
OBL	-0.042	0.390	-0.108	0.914
GEN	-3.509	1.403	-2.502	0.012*
DEF	0.783	0.213	3.672	<0.001*

Table 83: Frequency of Relative Clauses Referring to Inanimate Referents in IA Texts in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.019	0.497	0.038	0.970
ISU	0.683	0.716	0.954	0.340
DO	1.440	0.367	3.926	<0.001*
OBL	0.726	0.384	1.888	0.059.
GEN	-2.744	1.081	-2.538	0.011*
DEF	0.783	0.214	3.658	<0.001*

Table 84: Differences among IA Relative Clauses According to the Positions of their Heads in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.365	2.631	-1.659	0.097.
TSU	-1.325	3.996	-0.332	0.740
DO	3.966	2.604	1.523	0.128
OBL	-5.455	3.413	-1.599	0.110
GEN	-1.089	3.632	-2.998	0.003*
DEF	8.304	1.678	4.947	<0.001*
SU HEADS	-9.694	3.542	-2.737	0.006*
TOPIC HEADS	-5.448	2.627	0.000	1.000
OBL HEADS	-1.892	2.760	-0.686	0.493
GEN HEADS	-4.229	2.952	-1.432	0.152
SUCOM HEADS	4.169	2.392	1.743	0.081.
OTHERS HEADS	-8.824	3.433	-2.570	0.010*

Table 85: Frequency of Relative Clauses in Fiction Texts in Comparison with ISU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.934	0.154	6.051	<0.001*
TSU	-0.028	0.163	-0.170	0.865
DO	0.219	0.159	1.380	0.168
OBL	-0.354	0.179	-1.979	0.048*
GEN	-1.311	0.192	-6.842	<0.001*
DEF	0.973	0.060	16.167	<0.001*
INANIM	1.170	0.063	18.525	<0.001*
IA	-1.017	0.154	-6.600	<0.001*
MSA	0.190	0.11573	1.640	0.101

Table 86: Frequency of Relative Clauses in Fiction Texts in Comparison with OBL Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.546	0.234	2.330	0.020*
TSU	0.330	0.193	1.708	0.088.
DO	0.571	0.159	1.380	0.168
ISU	0.355	0.198	2.891	0.004*
GEN	-0.959	0.216	-4.442	<0.001*
DEF	0.973	0.060	16.167	<0.001*
INANIM	1.170	0.063	18.525	<0.001*
IA	-0.974	0.155	-6.274	<0.001*
MSA	0.253	0.117	2.154	0.131*

Table 87: Distribution of Relative Clauses in Fiction Texts while Considering Absolutive Relative Clauses as a Category

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.485	0.131	11.303	<0.001*
ABS	0.128	0.060	2.140	0.032*
DO	-0.460	0.100	-4.613	<0.001*
OBL	-1.053	0.158	-6.645	<0.001*
GEN	-2.014	0.171	-11.791	<0.001*
DEF	1.019	0.049	20.653	<0.001*
INANIM	1.203	0.052	23.266	<0.001*
IA	-0.888	0.132	-6.741	<0.001*
MSA	0.333	0.100	3.339	0.001*

Table 88: Distribution of Definite Relative Clauses in Fiction Texts in Comparison with ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.998	0.266	3.752	<0.001*
ISU	-0.126	0.258	-0.490	0.624
DO	0.279	0.090	3.109	0.002*
OBL	-0.411	0.290	-1.416	0.157
GEN	-1.725	0.240	-7.203	<0.001*
INANIM	1.270	0.076	16.682	<0.001*
IA	-0.181	0.355	-0.510	0.610
MSA	1.355	0.314	4.315	<0.001*

Table 89: Distribution of Definite Relative Clauses in Fiction Texts in Comparison to OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.623	0.457	1.364	0.173
ISU	0.298	0.155	1.916	0.055.
DO	0.707	0.280	2.529	0.011*
TSU	0.396	0.300	1.320	0.187
GEN	-1.280	0.216	-5.939	<0.001*
INANIM	1.270	0.076	16.682	<0.001*
IA	-0.139	0.359	-0.387	0.699
MSA	1.182	0.310	3.808	<0.001*

Table 90: Distribution of Indefinite Relative Clauses in Fiction Texts in Comparison to DO.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.565	0.253	6.175	<0.001*
TSU	0.101	0.163	0.617	0.537
ISU	-0.200	0.228	-0.879	0.379
OBL	-0.401	0.229	-1.747	0.081.
GEN	-0.844	0.277	-3.045	<0.001*
INANIM	0.927	0.114	8.136	<0.001*
IA	-1.528	0.270	-5.668	<0.001*
MSA	-0.675	0.223	-3.027	0.003*

Table 91: Distribution of Indefinite Relative Clauses in Fiction Texts in Comparison to OBL

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.155	0.241	4.800	<0.001*
TSU	0.464	0.164	2.825	0.005*
ISU	0.177	0.252	0.702	0.483
DO	0.371	0.216	1.719	0.086.
GEN	-0.474	0.307	-1.546	0.122
INANIM	0.927	0.114	8.138	<0.001*
IA	-1.495	0.273	-5.466	<0.001*
MSA	-0.610	0.226	-2.698	0.007*

Table 92: Distribution of Relative Clauses Referring to Inanimate Referents in Fiction Texts in Comparison with TSU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.770	0.176	10.080	<0.001*
ISU	0.130	0.224	0.580	0.562
DO	0.617	0.149	4.156	<0.001*
OBL	0.132	0.251	0.527	0.598
GEN	-1.325	0.196	-6.758	<0.001*
DEF	1.059	0.070	15.059	<0.001*
IA	-1.277	0.177	-7.231	<0.001*
MSA	0.240	0.123	1.954	0.051*

Table 93: Distribution of Relative Clauses Referring to Inanimate Referents in Fiction Texts in Comparison to ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.907	0.193	9.900	<0.001*
TSU	-0.129	0.231	-0.560	0.5756
DO	0.491	0.241	2.036	<0.042*
OBL	-0.015	0.273	-0.056	0.956
GEN	-1.462	0.251	-5.833	<0.001*
DEF	1.059	0.070	15.057	<0.001*
IA	-1.263	0.170	-7.433	<0.001*
MSA	0.213	0.119	1.788	0.074 .

Table 94: Distribution of Relative Clauses Referring to Animate Referents in Fiction Texts in Comparison to ISU

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.085	0.282	3.845	<0.001*
TSU	0.305	0.219	1.391	0.164
DO	-0.393	0.239	-1.641	0.101
OBL	-1.937	0.390	-4.972	<0.001*
GEN	-0.898	0.370	-2.426	0.015*
DEF	0.716	0.118	6.073	<0.001*
IA	-0.380	0.221	-1.721	0.085.
MSA	0.272	0.194	1.402	0.161

Table 95: Distribution of Relative Clauses Referring to Animate Referents in Fiction Texts in Comparison to GEN

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.277	0.225	1.230	0.219
ISU	0.898	0.301	2.888	0.004*
TSU	1.173	0.246	4.776	<0.001
DO	0.485	0.222	2.182	0.029
OBL	-1.064	0.383	-2.775	0.006*
DEF	0.716	0.118	6.083	<0.001*
IA	-0.456	0.241	-1.892	0.058.
MSA	0.171	0.209	0.815	0.415

Table 96: Differences among Fiction Relative Clauses According to the Positions of their Heads in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.179	0.152	7.741	<0.001*
TSU	-0.028	0.163	-0.169	0.866
DO	0.219	0.159	1.380	0.168
OBL	-0.354	0.179	-1.979	0.048*
GEN	-1.311	0.192	-6.842	<0.001*
DEF	0.973	0.060	16.166	<0.001*
SU HEADS	-1.445	0.112	-12.926	<0.001*
TOPIC HEADS	-1.165	0.100	-11.634	<0.001*
DO HEADS	-0.395	0.077	-5.130	<0.001*
GEN HEADS	-0.438	0.078	-5.619	<0.001*
SUCOM HEADS	-1.036	0.095	-10.860	<0.001*
OTHERS HEADS	-2.457	0.174	-14.137	<0.001*
IA	-1.017	0.154	-6.600	<0.001*
MSA	0.190	0.116	1.640	0.101

Table 97: Distribution of Relative Clauses in Nonfiction Texts

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.086	0.217	9.635	<0.001*
TSU	-0.819	0.359	-2.285	0.022*
DO	-0.335	0.168	-1.991	0.047*
OBL	-0.811	0.135	-6.025	<0.001*
GEN	-1.921	0.253	-7.588	<0.001*
DEF	1.001	0.061	16.503	<0.001*
INANIM	0.954	0.060	15.893	<0.001*
MSA	-0.744	0.154	-4.819	<0.001*

Table 98: Frequency of Relative Clauses in Nonfiction Texts in Comparison with DO Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.760	0.334	5.275	<0.001*
TSU	-0.487	0.239	-2.036	0.042*
ISU	0.333	0.169	1.969	0.049*
OBL	-0.478	0.122	-3.926	<0.001*
GEN	-1.588	0.221	-7.196	<0.001*
DEF	1.001	0.061	16.503	<0.001*
INANIM	0.954	0.060	15.893	<0.001*
MSA	-0.759	0.155	-4.881	<0.001*

Table 99: Frequency of Relative Clauses in Nonfiction Texts in Comparison with OBL Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.282	0.262	4.892	<0.001*
TSU	-0.009	0.323	-0.029	0.977
ISU	0.811	0.131	6.186	0.049*
DO	0.478	0.122	3.926	<0.001*
GEN	-1.111	0.192	-5.791	<0.001*
DEF	1.001	0.061	16.503	<0.001*
INANIM	0.954	0.060	15.893	<0.001*

Table 100: Distribution of Relative Clauses in Nonfiction Texts while Considering Absolutive Relative Clauses as a Category

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.219	0.200	11.113	<0.001*
ABS	0.112	0.067	1.673	0.094.
DO	-0.778	0.107	-7.281	<0.001*
OBL	-1.259	0.127	-9.951	<0.001*
GEN	-2.357	0.255	-9.243	<0.001*
DEF	0.996	0.048	20.739	<0.001*
INANIM	0.894	0.047	19.047	<0.001*
MSA	-0.036	0.210	-0.173	0.862

Table 101: Distribution of Definite Relative Clauses in Nonfiction Texts in Comparison with DO Relative Clauses.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.425	0.275	8.823	<0.001*
TSU	-0.701	0.316	-2.218	0.027*
ISU	0.171	0.175	0.978	0.0328
OBL	-0.473	0.161	-2.928	0.003*
GEN	-2.287	0.307	-7.455	<0.001*
INANIM	1.131	0.073	15.449	<0.001*
MSA	-0.189	0.130	-1.456	0.146

Table 102: Distribution of Definite Relative Clauses in nonfiction Texts in Comparison with TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.724	0.552	3.125	0.002*
ISU	0.872	0.430	2.028	0.043*
DO	0.701	0.316	2.218	0.027*
OBL	0.228	0.442	0.516	0.606
GEN	-1.586	0.364	-4.355	<0.001*
INANIM	1.130	0.073	15.449	<0.001*
MSA	-0.189	0.130	-1.456	0.146

Table 103: Distribution of Indefinite Relative Clauses in Nonfiction Texts in Comparison with TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.321	0.681	1.940	0.052.
ISU	0.878	0.401	2.193	0.028*
DO	-0.224	0.264	-0.848	0.397
OBL	-0.331	0.214	-1.547	0.122
GEN	-0.250	0.514	-0.485	0.628
INANIM	0.528	0.107	4.919	<0.001*
MSA	-0.915	0.400	-2.289	0.022*

Table 104: Distribution of Indefinite Relative Clauses in Nonfiction Texts in Comparison with OBL Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.296	0.771	1.680	0.093.
ISU	1.291	0.380	3.398	0.001*
TSU	0.186	0.284	0.653	0.514
DO	0.362	0.228	1.586	0.112
GEN	0.070	0.546	0.128	0.898
INANIM	0.528	0.107	4.920	<0.001*
MSA	-1.692	0.451	-3.751	<0.001*

Table 105: Distribution of Relative Clauses Referring to Animate Referents in Nonfiction Texts in Comparison with TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.948	0.388	5.023	<0.001*
ISU	0.883	0.188	4.687	0.001*
DO	-0.655	0.277	-2.364	0.018*
OBL	-1.736	0.328	-5.294	<0.001*
GEN	-1.416	0.402	-3.520	<0.001*
DEF	0.582	0.106	5.471	<0.001*
MSA	-1.027	0.342	-3.004	0.003*

Table 106: Distribution of Relative Clauses Referring to Animate Referents in Nonfiction Texts in Comparison with GEN Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.530	0.311	1.705	0.088.
ISU	2.302	0.333	4.687	<0.001*
TSU	1.418	0.403	3.519	<0.001*
DO	0.764	0.440	1.734	0.083.
OBL	-0.316	0.432	-0.732	0.464
DEF	0.582	0.106	5.471	<0.001*
MSA	-1.026	0.342	-3.004	0.003*

Table 107: Distribution of Relative Clauses Referring to Inanimate in Nonfiction Texts in Comparison with ISU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.347	0.209	11.255	<0.001*
TSU	-1.208	0.599	-2.015	0.044*
DO	0.120	0.140	0.857	0.391
OBL	-0.284	0.161	-1.765	0.078.
GEN	-1.717	0.260	-6.616	<0.001*
DEF	1.185	0.075	15.863	<0.001*
MSA	-0.188	0.164	-1.146	0.252

Table 108: Distribution of Relative Clauses Referring to Inanimate Referents in Nonfiction Texts in Comparison with TSU Relative Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.181	0.790	1.495	0.135
ISU	1.229	0.611	2.012	0.044*
DO	1.350	0.570	2.367	0.018*
OBL	0.948	0.672	1.410	0.158
GEN	-0.490	0.604	-0.811	0.417
DEF	1.185	0.075	15.861	<0.001*
MSA	-0.317	0.159	-1.993	0.046*

Table 109: Differences among Nonfiction Relative Clauses According to the Positions of their Heads in the Main Clauses

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.544	0.136	11.333	< 2e-16 *
TSU	-0.399	0.196	-2.033	0.042 *
DO	-0.075	0.151	-0.496	0.620
OBL	-0.551	0.137	-4.011	<0.001*
GEN	-1.536	0.161	-9.552	<0.001*
DO HEADS	-0.381	0.057	-6.734	<0.001*
GEN HEADS	-0.406	0.057	-7.123	<0.001*
OTHERS HEADS	-2.569	0.135	-19.010	<0.001*
SUCOM HEADS	-1.267	0.077	-16.473	<0.001*
SU HEADS	-1.166	0.074	-15.763	<0.001*
TOPIC HEADS	-1.208	0.075	-16.074	<0.001*
DEF	0.934	0.044	21.135	<0.001*
IA	-1.309	0.218	-6.004	<0.001*
MSA	0.031	0.161	0.193	0.847

Appendix C: Classification of relative clauses according to the grammatical positions of the head NPs in the main clauses and in the relative clauses

- **SU(TOPIC)- ISU:**

1. inna al-muslim-īn allaḏīn ʿabar-ū ʾila
   indeed the-muslim-PL REL(3.MS.PL) cross-PL to
   faransa kan-ū bi-aʾʿdād kabīrat
   France was-PL in-numbers big
   Indeed, Muslims who went to France were many.
   [MSA, 4: 57]

- **SU(TOPIC)-TSU**

2. inna ʾawal malik malaka fī al-ʾarḍ šarqa-ha
   indeed first king governed in the-earth east-its
   wa ġarb-ha namrūd
   and west-its Namroud
   Indeed the first king who governed the east and west was Namroud.
   [CA, 6:81]

- **SU(TOPIC)-DO**

3. al-šaiʾ al-waḥīd allaḏī raġibt-ah min
   the-thing the-sole REL(3.SG.MAS) want.I-it from
   hāḏā al-qāʾid huwa al-ṣamt
   this the-commander it the-silance
   The only thing which I want from this commander is to keep silent.
   [MSA, 3: 48]

- **SU(TOPIC)-IO**

4. kān šeiḫ al-qabīlah allaḏī quddim-
   was sheiḫ the-tribe REL(3.SG.MAS) presented-
   at la-hu al-hadiyah yataġāḍā ʿan hāḏā al-ǧawr
   FEM to-him the-present overlook about this the-injustice

The sheikh of the tribe to whom the present is given overlooked this injustice.
[MSA, 3: 174]

- **SU(TOPIC)-OBL**

5. al-bait allaḏī huwa sākin fī-hi
   the-house REL(3.SG.MAS) he resident in-it
   baitu-nā
   house-our
   The house in which he lives is our house.
   [CA, 1: 311]

- **SU(TOPIC)-OCOMP**

6. inna ġazāl lam yurā ġazāl aḥsan
   indeed deer not seen deer better
   min-hu ǧīʾ bi-hi
   from-it brought in-it
   Indeed, a deer which there was no deer better than was brought.
   [CA, 4: 749]

- **SU-ISU**

7. ṣaʿad al-raǧul allaḏī kān ʾila ǧānib-ha
   went up the-man REL(3.SG.MAS) was to side-her
   The man who was beside her went up.
   [MSA, 1: 34]

- **SU-TSU**

8. ṣadar-at al-tašrīʿ-āt allatī
   issued-FEM the-legislations-FEM REL(3.SG.FEM)
   tastahduf ḥimayat al-muwaṭin al-ʿomani
   target protection the-citizen the-Omani
   The legislations that target the protection of the Omani citizen were issued.
   [MSA, 6: 208]

- **SU-DO**

9. falma daḫala-t al-sanah allatī waṣaf
   when entered-FEM the-year REL(3.SG.FEM) described
   aṣḥāb al-nuğūm li-namrūd...
   owners the-stars to-Namroud...
   When the year which the astrologers described to Namroud started...
   [CA, 4: 81]

- **SU-OBL**

10. min ʾain ğāʾa-t tilka al-makānah allatī
    from where came-FEM that the-status REL(3.SG.FEM)
    waḍaʿ-ū-h fī-ha
    put-they-him in-it
    From where the status in which they put him came?
    [MSA, 5: 115]

- **SU-GEN**

11. daḫal šakṣ qad aḥnā al-ḍahr
    entered person indeed bowed the-age
    saʿdat-ah
    stature-his
    A man whose stature was bowed by age entered.
    [CA, 3: 141]

- **SU-OCOMP**

12. fī-hum ġulām lais fī al-ġulmān ġulām
    in-them boy no in the-boys boy
    aḍraf min-hu
    nicer than-him
    Among them there is a boy whom there is no one nicer than him.
    [CA, 4: 127]

- **DO-ISU**

13. naqaḍa al-rūm al-ṣulḥ allaḏī kān
    repeal the-Roman the-reconcilation REL(3.MAS.FEM) was
    bain-hum wa bain al-muslimīn
    between-them and between the-Muslims
    Roman repealed the reconciliation which was with Muslims
    [CA, 5: 852]

- **DO-TSU**

14. wa istadʿai-tu al-malāḥ allaḏī aḫrağa-nā
    and summoned-I the-sailor REL(3.SG.MAS) take out-us
    tilk al-lailah
    that the-night
    And I summoned the sailor who take us out that night.
    [CA,1:81]

- **DO-DO**

15. rafaḍ siğāra-tī allatī qaddam-tu-hā ila-hi
    refused cigarette-my REL(3.SG.FEM) presented-I-it to-him
    He refused my cigarette that I have presented to him.
    [MSA, 2: 10]

- **DO-OBL**

16. ma ʾa-ḥib al-banāt yšof-ūn ḏīch al-makān-āt
    not I-like the-girls see-PL those the-locality-PL
    wa al-makān al-daʿīš bi-ha
    and the-place REL-live in-it
    I do not like that the girls see those localities and place in which I live.
    [IA, 1: 77]

- **DO-GEN**

17. iğtamaʿ-ū wa qadam-ū ʿala anfus-i-hum rağul
gathered-PL and presented-PL on selves-GEN-them man
ʾismu-hu muhammed bin ʿisa
name-his Muhammed son Isa
The gathered and chose a man whose name is Muhammed- the son of ʿisa their leader. [CA, 5:982]

- **DO-OCOMP**

18. raʾi-tu li-ḏalik al-malik bint lam yuḫlaq
saw-I to-that the-king daughter not creat.PASS
fī zamāni-hā aḥsan min-hā
in time-her better from-her
I saw a king's daughter whom no one at that time was better than.
[CA, 1: 49]

- **IO-ISU**

19. iʿtāʾa baʿḍ al-durūs al-kuṣūṣiyah bi-l-luġah al-
Giving some the-lesson the-tutorial in-the-language the-
-innğiliziyah li-saiydah kāna-t taʿmal fī ʾiḥdā
-English to-lady was-FEM work in one
ḥadāʾiq al-safārah
gardens the-embassies
Giving some English tutorial lessons to a lady who was working in one of the embassies' gardens. [MSA, 1: 76]

- **IO-GEN**

20. arsal al-hadāya ʾila wazīr al-malik
sent the-gifts to minister the-king
allaḏī tazawağ bi-hā ibnu-hu
REL(3.SG.MAS) married from-her son-his
He sent the gifts to the king's minister whose son has got married from her.
[CA, 1: 180]

- **OBL-ISU**

21. ka-wāḥid min haḏ-ih al-aǧsām allatī la
like-one of these-FEM the-bodies REL(3.SG.FEM) not
taḥis wa la tataġaḏā
feeland not eat
As one of these bodies that do not feel and do not eat. [CA, 2: 48]

- **OBL-TSU**

22. wa awarad ʿadad min al-riwaya-āt
and cited number of the-narration-PL
allatī tubṭil qawla-hu
REL(3.SG.FEM) invalidate saying-his
And he cited a number of narrations that invalidate his sayings.
[MSA, 5: 23]

- **OBL-DO**

23. ṯum qāl fī al-maktūb allaḏī kataba-hu
then stated in the-letter REL[3.SG.MAS] wrote-it
li-wālid-i
to-father-my
Then (he) stated in the letter which he wrote to my father...
[CA, 1: 221]

- **OBL- OBL**

24. wa ta-ḥmilu-hu ʾila mawādiʿ fī-hā
and FEM-carried-him to places in-them
šaǧar muṯmir
trees fruitful
And (she) carried him to places in which there are fruitful trees.
[CA, 2: 27]

- **OBL-GEN**

25. li-yakūn ḥadīṯ al-nās fī muǧtamaʿ lam
to-be talk the-people in society not
tataṭawar ṯaqāfatu-hu wa mafāhīmu-hu
develop cultur-its and concepts-its
To be people's talk in a society whose culture and concepts have not been developed yet.
[MSA, 3: 137]

- **GEN-ISU**

26. kān ʾism aḥad alfta-yeiyn allaḏein udḫil-
was name one boy-DU REL(3.DU.MAS) enter.PASS-
-laa al-siǧn Maḥlab
-DU the-prison Mahlab
The name of one of the boys who are put into prison was Mahlab.
[CA, 4: 116]

- **GEN-TSU**

27. kān ʾistqrār al-barbr allaḏīn daḫal-ū
was settlement the-Barbar REL(3.PL.MAS) enter-PL
al-ʾandalus...mumāṯil li-istqrār al-baladīn
the-Andalus...similar to-settlement the-Baladīn
The settlement of the Barbar who entered al-Andalus was similar to that of the Baladīn.
[MSA, 4: 80]

- **GEN-DO**

28. iḏāʿet tafāṣīl al-itifāq-āt allatī
announcement details the-agreement-PL REL(3.SG.FEM)
waqaʿʿat-hā Russia
signed-them Russia
The announcement of the details of the agreements which Russia signed.
[MSA, 6: 57]

- **GEN-IO**

29. kān ʾawal wāli faraḍ la-hu raʿiyatu-hu
    was first governor assigned to-him people-his
    nafaqāt-hu
    salary-his
    He was the first governor to whom his citizens assigned a salary.
    [CA, 5: 306]

- **GEN-OBL**

30. an yakūn li-l-šamyīn ṯulṯ-ā al-ʾaraḍī wa
    to be to-the-Shamyin third-DU the-land and
    al-mazāriʿ allatti yaskun-ūn fī-hā
    the-farms REL(3.SG.FEM) live-PL in-them
    Two Thirds of the land and the farms in which the Shamieen live are to be for them.
    [MSA, 4: 76]

- **GEN-GEN**

31. ...mawqif al-quwwat-ein al-ʾurubiat-ein
    ...position the-power-DU the-European-DU
    allatein istawla-t quwwa-āt-huma ʿalā al-arāḍī
    REL(3.DU.FEM) seize-FEM power-PL-their on the-lands
    al-lībiyah
    the-Libyan
    ...The position of the two European countries whose powers seized the Libyan lands
    [MSA, 6: 120]

- **OCOMP-ISU**

32. le-ǧawāb yušfī ẖair min ǧilbāb yuddaffī
    for-answer heal better than dress warm
    A sufficient answer is better than a dress that keeps (you) warm.
    [CA, 3: 255]

- **OCOMP-TSU**

33. inna al-wāqiʿ... ašad ʿunfan min kul al-
indeed the-reality... more violent of all the-
-ẓurūf alsaiʾah allatī taġzū hayāta-nā
- circumstances bad REL(3.SG.FEM) invade life-our
Indeed the reality is ... more violent than all the bad circumstances that invade our life.
[MSA, 3: 24]

- **OCOMP-DO**

34. ʾu-ʿṭī-hi akṯar min māla-hu allaḏī
I-give-him more of money-his REL(3.SG.MAS)
aḫaḏ-tu-hu min-hu
took-I-it from-him
I give him money more than the money which I took from him.
[CA, 1: 325]

- **OCOMP-OBL**

35. u-ġādir mahzūm ẖeir min ʾintiṣār yakūn fī- hi
I-leave defeated better of victory is in- it
šaqiq-i huwa al-mahzūm
brother-my he the-defeated
For me, to be defeated is better than a victory in which my brother is the one who is defeated.
[MSA, 3: 23]

- **SUCOM-ISU**

36. wa huwa al-kawkab al-samāwī allaḏī
and he the-planet the-sky REL(3.SG.MAS)
yadūr ḥawla-hā
spin around-her
And he is the sky planet that goes around her. [MSA, 3: 10]

- **SUCOM-TSU**

37. hāḏā ḥachi mā ywakil ḫubuz
    this talk not feed bread
    This is talk that does not feed bread.
    [IA, 2: 38]

- **SUCOM-DO**

38. hai yimkin falsafah mā a-ʿruf-hā zīn
    this may be philosophy not I-know-it well
    This may be a philosophy which I do not know well. [IA, 1: 60]

- **SUCOM-OBL**

39. wa hia ṯaman al-ṭaʿām allaḏi
    and it price the-food REL(3.SG.MAS)
    ištar-ū-hu bi- hi
    bought-they-it in-it
    It is the price by which they bought the food. [CA, 4: 118]

- **SUCOM-GEN**

40. kaʾan-nī fatāt iktašaf abū-hā faǧʾatan inna-hā
    as if-I girl discovered father-her suddenly indeed-she
    dāʿirah
    prostitute
    As if I was a girl whose father discovered suddenly that she was a prostitute.
    [MSA, 2: 83]

- **SUCOM-OCOMP**

41. hia al-kifāyah allatti lā baqāʾ
    it the-sufficiency REL(3.SG.FEM) not survival
    li-l-rūḥ al-ḥaiwanī bi-ʾaqal min-hā
    for-the-spirit the-living by-less of-it
    It is the sufficiency by less than which there is no survival for living beings.
    [CA, 2: 85]

- **EXIST-ISU**

42. ma-kū šī yfūt wa mā tḥis bī-h
no-there thing go and no feel in-it
There is nothing that goes without you feeling it.
[IA, 1: 86]

- **EXIST-TSU**

43. hunāk riwāyāt uẖrā tušīr ʾilā ʾinna-hu ʾistmar
there narration other refer to indeed-he continued
fī futūḥāti-hi
in conquests-his
There were narrations which indicate that he continued in his conquests.
[MSA, 4: 37]

- **EXIST-DO**

44. kān hunāk niẓām ǧadīd lam ʾu-šāhid-hu min
Was there system new not I-see-it of
qabl
before
There was a new system which I have not seen before.
[MSA, 3: 36]

- **EXIST-OBL**

45. ʾinna hunāk waṣyiah yuḥāwil al-baʿḍ al-taʿtīm
indeed there will try the-some the-covering
ʿalai-hā
on-it
Indeed, there is a will on which some people are trying to put a cover.
[MSA, 5: 44].

- **EXIST-GEN**

46. mā-kū waḥd-at abū-ha ʿaṭṭār
    no-there one-FEM father-her grocer
    Is not there anyone whose father is a grocer?
    [IA, 2: 91]

- **CACC-TSU**

47. hadm ay muḥāwalah li-tafsīr al-naṣ
    destroy any attempt to-interpret the-text
    tafsīran yamis al- ṣahabah
    interpretation touch the-sahabah
    (They) reject any attempt to give the texts (of Qurʾan) an interpretation that affects the sahabah.
    [MSA, 5: 28]

- **CACC-DO**

48. la-ya-ḍriban-kum ḍarban yartāb bi-hi al-mubṭil-
    will-he-hit-you.PL hitting suspect in-it the-bad-
    -ūn
    -PL
    He will hit you hitting which the bad people suspect.
    [CA, 4: 855]

- **CACC-OBL**

49. fa-laṭam waǧha-hā laṭmatan ʾiḥḍarr-at
    then-slapped face-her slap became.green-FEM
    ʿain-hā min-hā
    eye-her of-it
    Then he slapped her face a slap because of which her eye became green.
    [CA, 4: 422]

- **CACC-GEN**

50. wa ḥazan ʿalai-hi huzn lam yusmaʿ bi-miṯli-hi
    and grieve on-him grief not hear.PASS in-similar-it
    and (he) grieved for him a grief whose no similar one.
    [CA, 6: 899]

- **ADV-DO**

51. kaif yabqā hāḏā al-buḫār al-muddah alattī
    how stay this the-steam the-period REL(3.SG.FEM)
    baqā
    stay
    How does steam stay the while which it stays?
    [CA, 2: 40]

- **ADV-OBL**

52. fa-nāǧaza-hum saʿat tallaqā-hum
    then-fought-them hour met-them
    Then (he) fought them in the time he met them.
    [CA, 4: 1103]

- **ADV-GEN**

53. fa-qaḍay-nā-hā lailah ġābat šawāʾibu-hā
    so-spent-we-it night disappeared impurities-its
    So we spent a night whose impurities disappeared [a good night]
    [CA, 3: 55]

- **OCOM-DO**

54. fa-waǧada-hu insān lam yaʿrif-hu
    then-found-him human being not know-him
    Then he found him as a human being whom he does not know.
    [CA, 1: 78]